Characterisation of structured surfaces and assessment of associated measurement uncertainty

Gavin MacAulay

October 26, 2015

Abstract

Recently, structured surfaces, consisting of deterministic features designed to produce a particular effect, have shown promise in providing superior functional performance for a range of applications including: low friction surfaces, hydrophobic surfaces and optical effects.

Methods have been developed to characterise such structured surfaces. The most widely used characterisation methods are based on segmenting the surface in feature and background regions and then determining the geometrical properties of those features. However, further work is needed to refine these characterisation techniques and provide associated uncertainties.

This thesis considers the effect of various segmentation control parameters such as thresholds on the final geometric parameters. The effect of varying filter size is also considered. These considerations should help in selecting a suitable characterisation method for future projects.

Additionally, uncertainty in the characterisation should be estimated in order to give an indication of the accuracy of the assessment. However, no previous work has assessed uncertainty in the dimensional properties of structured surfaces. Therefore, this thesis presents two methods to characterise the uncertainty in the geometric characteristics of structured surfaces.

First, the measurement reproducibility is used, which can be determined by repeated measurement of a feature. However, measurement reproducibility cannot account for all sources of uncertainty and cannot assess any bias in the measurements.

Therefore, a second method based on assessment of the metrological characteristics of the instrument is considered. The metrological characteristics estimate errors produced by the instrument in a way that can easily be measured. Monte Carlo techniques are then used to propagate the effects of the metrological characteristics and their uncertainties into the final measurement uncertainty. For the example used, it was found that the results using the metrological characteristics were in good agreement with the reproducibility

results.

From these results, it is concluded that the choice of segmentation method, control parameters and filtering can all significantly effect the characterisation of features on a structured surface, often in unexpected ways. Therefore, care must be taken when selecting these values for a specific application. Additionally, two methods of determining the uncertainty of the structured surfaces were considered. Both methods are valid and produce similar results. Using the measurement reproducibility is simple to perform, but requires many measurements and cannot account for some uncertainty sources such as those due to the instrument amplification factors. On the other hand, the use of metrological characteristics can account for all significant sources of uncertainty in a measurement, but is mathematically more complex, requiring Monte Carlo simulations to propagate the uncertainties into the final characteristics. Additionally, other artefacts than the sample being measured are required to determine the metrological characteristics, which may be an issue in some cases.

Acknowledgements

I would like to thank all the people who have helped me throughout my PhD. In particular my supervisors: Prof Richard Leach, Dr Claudiu Giusca and Dr Atanas Ivanov whose advise and assistance have been essential throughout the project. Without their support this project would not have been possible.

Additionally, thanks to all my colleges at NPL and beyond. Particularly, Nicola Senin, Andrew Henning, James Claverley, Daniel O'Conner and Chris Jones. Without your friendship and support this project would have been more difficult and less enjoyable.

I am grateful for the support of Brunel University and the National Physical laboratory who have funded this work.

Finally, thanks to my parents without whom I would never have reached this stage.

Contents

List of Figures

List of Tables

List of abbreviations

ACF Autocorrelation function

AFM Atomic force microscope

GUM Guide to the expression of uncertainty in measurement

ICP Iterative closest point algorithm

ISO International standards organisation

NA Numerical aperture

OOR Out of roundness

STM Scanning tunnelling microscope

VIM International vocabulary of metrology

List of symbols

α Coefficient in geometric active contour algorithm.

α_z Vertical amplification coefficient.

β Coefficient in geometric active contour algorithm.

ϵ Estimate of the noise associated with a pixel.

λ Coefficient in geometric active contour algorithm.

λ_1 Weighting coefficient in active contours without edges algorithm.

λ_2 Weighting coefficient in active contours without edges algorithm.

λ_c Cut-off wavelength of a Gaussian filter.

μ Mean of measurements/values. Also used as smoothing coefficient in active contour without edges algorithm.

ν Coefficient in active contours without edges algorithm.

σ Standard deviation, sometimes abbreviated by std.

χ^2_{n-1} Chi squared distribution with $n-1$ degrees of freedom.

ϕ Diameter of a dimple.

Σ Co-variance matrix.

a Half the width of a triangular or uniform distribution, such that if centre on 0 the bounds are at $\pm a$.

c Sensitivity coeffcient.

c_1 Mean height inside contour in active contours without edges algorithm.

c_2 Mean height outside contour in active contours without edges algorithm.

C The contour position in the active contour algorithm.

CTrap$(-a, a, d)$ A curvilinear trapezoid centred on zero with nominal bounds at $\pm a$, with uncertainty in those bounds of $\pm d$.

d Depth of a dimple.

E Energy of the active contour.

F An estimate of the sharpness.

I Intensity of light.

k Coverage factor of a confidence interval, i.e. the number of standard deviations covered by the interval. Also the number of clusters in k-means clustering.

n Number of measurements/points/samples etc., normally a subset of the entire population.

N Number of measurements/points/samples etc., where N normally describes the entire population.

$N(\mu, \sigma)$ A normal distribution with mean μ and standard deviation σ.

P Probability density of a function.

r Correlation coefficient.

R Ratio between of measurement to manufacturing reproducibility, given as a percentage.

s Estimated standard deviation.

Sa Arithmetic mean height of a surface.

Sq Root mean square height of a surface.

Ssk Skewness of a surface.

Sku Kurtosis of a surface.

Sp Highest peak height.

Sv Lowest valley height.

Sz Maximum height of the surface, $Sp + Sv$.

Sal Autocorrelation length.

Str Texture aspect ratio.

Sdq Root mean square gradient of a surface.

Sdr Developed area ratio of a surface.

$Smr(c)$ Areal material ratio of a surface.

Sk Core height of a surface.

Svk Reduced peak height of a surface.

Spd Density of peaks

$S10z$ Average of 10 highest and lowest peaks, $S5p + S5v$.

$S5p$ Average of the 5 highest peaks.

$S5v$ Average of the 5 deepest valleys.

Sda Average dale area.

T Threshold value.

u An estimate of uncertainty.

u_c An estimate of the combined uncertainty, due to several different sources.

$U(a, b)$ A uniform distribution with lower bound a and upper bound b.

v Number of degrees of freedom.

z Height of a pixel/point.

Chapter 1

Introduction

1.1 Background

Structured surfaces, defined as surfaces consisting of a deterministic pattern of geometric features designed to produce a particular function (Evans and Bryan 1999; Stout and Blunt 2001), are a new class of surface that has become increasingly popular in recent years. They have shown promise in improving surface performance in a variety of applications including: tribological performance (Etsion et al. 1999; Wang 2014), super-hydrophobicity (Martines et al. 2005), cell adhesion (Yang et al. 2009), metamaterials (Pendry et al. 2004; Leskova et al. 2007) and anti-reflectance (Stavroulakis et al. 2013).

One of the limiting factors effecting the mass adoption of structured surfaces in real applications is the lack of appropriate metrology tools to analyse them. Without appropriate metrology it is difficult to correlate feature size with functional performance or provide suitable tolerances. This lack of suitable measurements can then increase cost, due to overly tight tolerances, or decrease performance, due to wide tolerances or an unsuitable choice of parameters.

Due to the micro/nano scale of individual features and their large spatial extent, consisting of many thousands of features across a surface, structured surfaces are generally considered as part of the surface texture from a characterisation point of view. However, techniques commonly used when characterising stochastic, non-structured, surfaces have significant drawbacks when characterising structured surfaces and may be ineffective. Recently, several researchers have considered an alternative characterisation approach based on segmenting the surface to identify the features and characterising geometrical

properties of those features in a similar way to how conventionally sized components are characterised (Zhu 2012; Blunt and Xiao 2011; Hartmann and Loderer 2014; Senin et al. 2012a).

This approach of segmenting the surface and characterising geometric properties of structured surfaces has been shown to be successful in identifying relevant parameters. Therefore, this thesis looks to further develop such approaches. In particular, while many of the characterisation techniques used for structured surfaces are developed from fields such as machine vision and image processing, how they are applied and what factors are significant when applying these techniques to structured structures need to be considered further.

Therefore, it is important to consider factors such as how choice of segmentation algorithm and controlling parameters affect the segmentation and resulting geometric properties. These choices are of critical importance as they define how the feature is identified. Up until now only cursory consideration has been given in the literature to how these choices affect the results.

Additionally, measurement uncertainty is an important factor to consider. While traditional surface texture parameters are often presented without uncertainty, it is at least possible to calculate the uncertainty associated with a measurement if desired (Haitjema 2013). However, for structured surfaces, no method has previously been published to determine the measurement uncertainty. This is an important effect which should be considered.

1.2 Aims and objectives

The work in this thesis aims to further develop techniques to characterise structured surfaces by considering how the choice of algorithm and various algorithm control parameters affect the measurement result. Additionally, it aims to develop techniques to assess the uncertainty in the measurement of structured surfaces. In particular the objectives of this thesis are to:

- Review the currently used approaches and provide a coherent approach to the characterisation of structured surfaces.
- Implement specific algorithms to allow for the characterisation of structured surfaces.

- Investigate the effect of changing segmentation algorithm and control parameters on the geometric properties of a surface.
- Develop and apply methods to assess the uncertainty of the characterisation of structured surfaces.
- Consider how different approaches and assumptions affect the final uncertainty and consider the validity of these choices.

1.3 Outline of thesis

The rest of the thesis is organised as follows. Chapter 2 provides a review of the relevant literature, including a definition of structured surfaces, characterisation methods for structured surfaces, and assessment of measurement uncertainty. Chapter 3 describes the samples, measuring instruments and characterisation algorithms used in the rest of the thesis. Chapter 4 investigates the effect of changing the characterisation method, the control parameters of these methods and the filtering used on the geometric properties of a feature. Chapter 5 presents an initial assessment of measurement uncertainty via repeated measurement of a feature. Chapter 6 investigates an alternative method to assess measurement uncertainty using the metrological characteristics of the measuring instrument and compares this approach to that in chapter 5. Chapter 7 summarises the results presented in this thesis, showing the main contributions and discussing possible directions for future work. Figure 1.1 summarises the areas covered by the thesis and shows how they are interrelated. The main focus of the thesis is on developing methods to determine uncertainty in parameters for measurement of structured surfaces and two methods to do this are show in chapters 5 and 6. However, the effect of varying segmentation method is very important as it influences both the parameters themselves and their uncertainty.

1.4 Key contributions

The main contributions of this thesis are highlighted below:

- A thorough review of current surface segmentation techniques for structured surfaces has been carried out. This work provides a coherent framework for the analysis of structured surfaces against which new methods can be compared and identifies areas in which further research is needed.

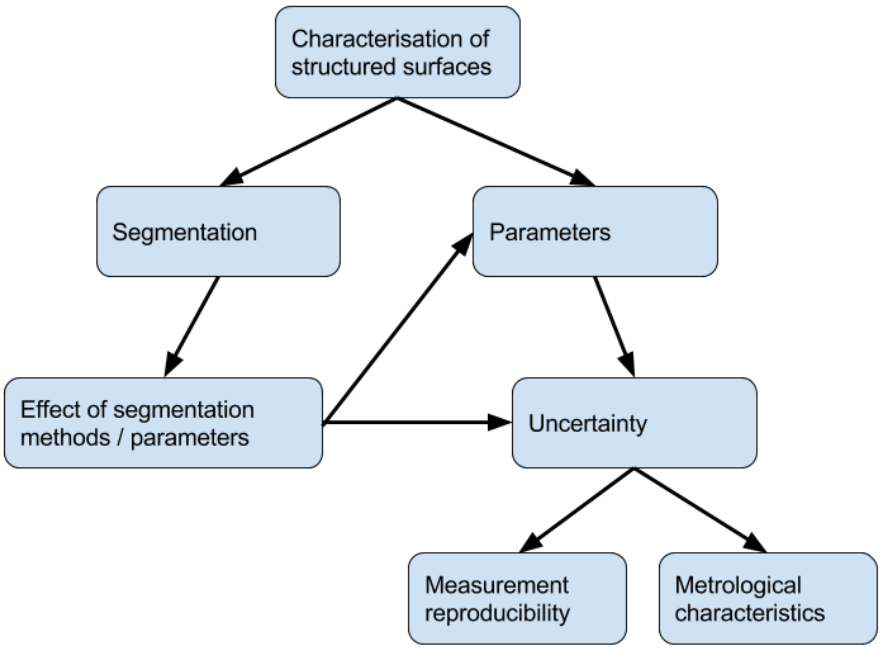

Figure 1.1: Overview of the areas dealt with by this thesis and how they are interconnected.

- Demonstrated the effects of various segmentation algorithms and parameters on structured surfaces. The range of results achieved by these different approaches mean there can be no unique definition of a feature.
- Investigated the effect of varying filter size on structured surfaces. It was found that there is no simple relationship between filter size and measured parameters.
- The measurement and manufacturing reproducibility of the features was investigated in a way that could easily be adapted for other surfaces.
- The approach to determine reproducibility was also used to investigate position dependence of the measurements.
- A method to determine measurement uncertainty using Monte Carlo simulations was developed. This approach provides a robust method to determine uncertainty using minimal additional measurements of the surface.

Chapter 2

Review of current literature

This chapter provides a review of the current literature relating to the characterisation of structured surfaces. First, the definition and applications of structured surfaces are defined and discussed. Next, surface metrology for conventional surfaces is considered, including the instrumentation and parameters used. From this starting point, the characterisation of structured surfaces is considered, including possible methods and the differences between characterising conventional and structured surfaces. Characterisation methods based on segmenting the features from surrounding topography are found to show promise for characterising structured surfaces and so are considered in further depth. The segmentation step, where the significant features are separated from the background is the key step in such characterisation approaches. Therefore, a number of segmentation methods that have previously been applied to the characterisation of structured surfaces are analysed and discussed. Finally, the principle of measurement uncertainty and its application to structured surfaces is reviewed.

2.1 Definition of structured surfaces

Evans and Bryan 1999 provide one of the first definitions of structured surfaces as, "surfaces with a deterministic pattern of usually high aspect ratio geometric features designed to give a specific function". Subsequent researchers have added refinements and further subdivisions. Stout and Blunt 2001 subdivided structured surfaces into directional and non-directional surfaces depending on whether the pattern has a dominant direction. Similarly, Jiang and Whitehouse 2012 subdivided structured surfaces into additional subcategories: tessellated, linear, rotationally invariant and multi-patterned as

shown in figure 2.1. Tessellated surfaces are structured surfaces with a periodic pattern. The most common subclass of tessellated surface has tiles with translational symmetry (e.g. retro-reflector surfaces consisting of a Cartesian pattern of hexagonal prisms, or the dimpled surface of a golf ball). The other subclasses of tessellated surface are linear patterned surfaces consisting of a single linear texture unit, which is repeated in only one direction (e.g. ribbed or grooved structures), and rotationally invariant patterned surfaces, where periodicity is in a polar coordinate system rather than a Cartesian one (e.g. Fresnel lenses). Finally, structured surfaces which do not have a periodic pattern are defined as multi-patterned surfaces, an example being many MEMS devices. This thesis focuses on tessellated surfaces and does not deal directly with other types of surface. However, many of the approaches and techniques discussed could be adapted for the analysis any type of structured surface with minor adjustment.

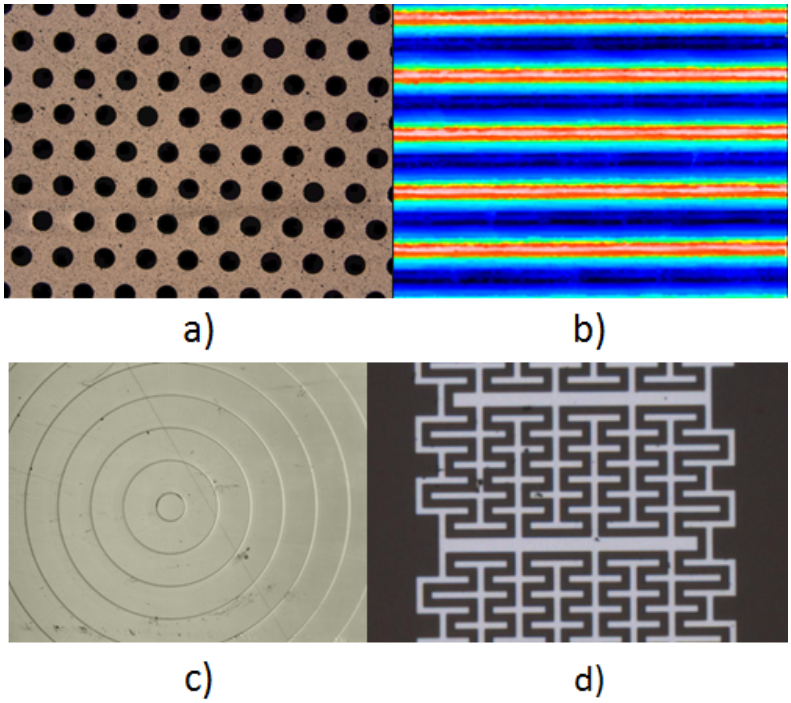

Figure 2.1: Examples of different types of structured surfaces: a) A tessellated surface with a translational pattern, a laser textured silicon nitride disk. b) A linear patterned surface, a micro-milled steel riblet surface. c) A rotationally symmetric Fresnel lens surface. d) A multi-patterned surface, part of a silicon quantum oscillator chip.

As shown in figure 2.2, structured surfaces are made of topologically connected pattern units, or tiles. Each tile consists of the same nominal topography, repeated multiple

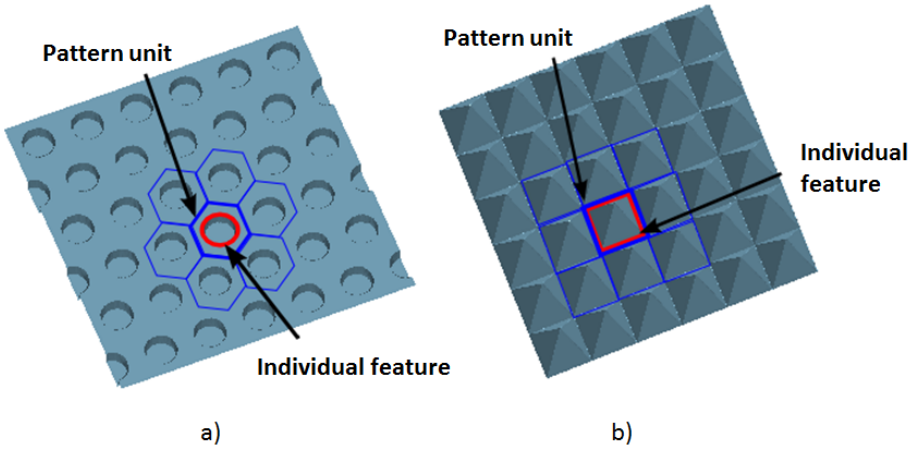

Figure 2.2: Pattern units or tiles (blue) and individual features (red) of two different model surfaces. a) A dimpled surface where the features only fill part of the tile and are surrounded by background surface. b) A pyramidal surface where the feature fills the entire surface and no background exists.

times to form the periodic pattern. The tiles are occupied by the deterministic features which make up the structured surface (e.g. the dimple of a dimpled pattern, a polygonal prism in retro-reflectors or abrasive surfaces). These features may cover all or part of the pattern unit. When the functionally relevant feature does not occupy the entire tile, the space between adjacent features can be referred to as the background. The distinction between feature and background is somewhat arbitrary, since the background surface is still important to the function. However, this distinction is often adopted when characterising structured surfaces as it resembles the original surface design specifications.

It is important to note that all structured surfaces will inevitably have a stochastic as well as deterministic component. The relative size of these components is a major factor which determines whether a surface is viewed as structured or not. However, there is a significant cross over, such as for stratified surfaces which have both significant stochastic and deterministic components (Jiang and Whitehouse 2012). This thesis does not consider this cross-over and will only consider the deterministic component of surfaces.

2.2 Applications of structured surfaces

In recent years, structured surfaces have grown rapidly in popularity and are now used in a wide variety of applications. Evans and Bryan 1999, Bruzzone et al. 2008, and Malshe et

al. 2013 have produced reviews describing the range of applications of structured surfaces. Some of the main applications are summarised here.

A popular application for structured surfaces is friction reduction in sliding surfaces, such as bearings, examples of which are shown in figure 2.3. Structured surfaces for this application have been widely investigated (Etsion et al. 1999; Ryk et al. 2005; Uehara et al. 2004; Andersson et al. 2007; Fowell et al. 2012; Wang et al. 2003) and results indicate that significant reduction in friction is possible. These surfaces generally consist of a repeated pattern of micro-scale pits; which produce lift, increasing separation between the surfaces and, therefore, reducing friction. A similar category of structured surface are riblets (Nieuwstadt et al. 1993; Bechert et al. 2000). These are designed for reduction of fluid drag, for example over aircraft wings and consist of high aspect ratio ridges perpendicular to the flow direction, which control the formation of the boundary layer.

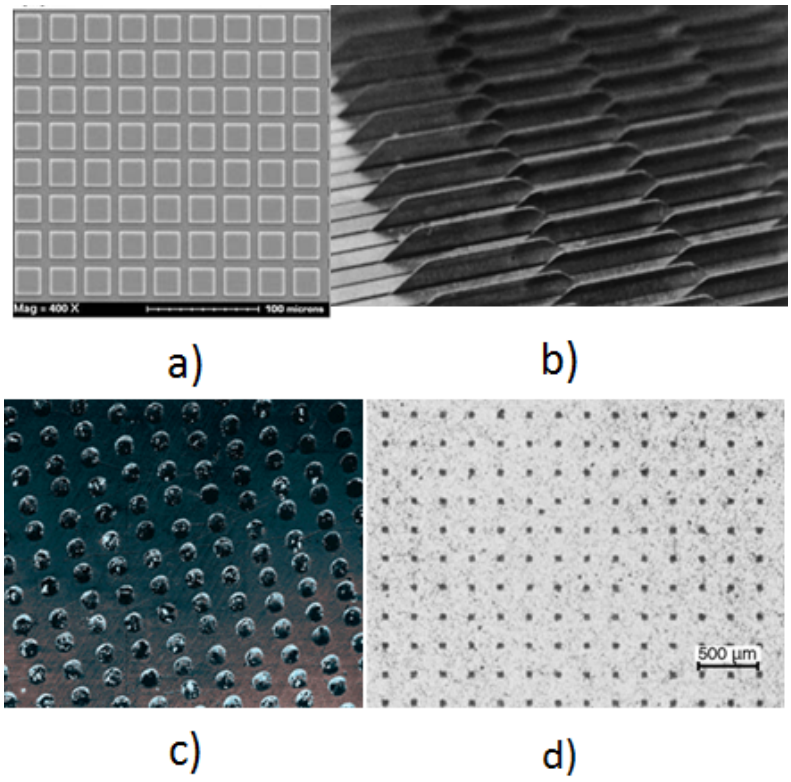

Figure 2.3: Example structured surfaces for friction reduction: a) Etched squares on a silicon surface (He et al. 2008), b) Staggered riblet structures (Bechert et al. 2000) c) Laser textured steel surface (Etsion 2005), d) Textured silicon carbide (Wang et al. 2006).

There are many other applications of structured surfaces, some of which are shown in figure 2.4. One application is optical effects, and many optical components, such as

Fresnel lenses or micro-lens arrays (Evans and Bryan 1999) are a form of structured surface. Similarly, structuring can alter the reflectance properties of surfaces, the most famous example of which is the moth eye effect (Stavroulakis et al. 2013; Evans and Bryan 1999). Such surfaces consist of high aspect ratio, sub-wavelength pillars that mimic the surface of a moth's eye and produce a low reflectance surface.

Similarly, structured surfaces have many biomedical applications, primarily for the control of cell adhesion and cell growth. Yang et al. (Yang et al. 2009) have shown that adding narrow grooves can encourage cell growth on a surface. Similarly, Crouch et al. (Crouch et al. 2009) have shown that varying groove width can control alignment and elongation of cells growing on the surface. Although, in many cases a combination of random and structured surfaces are used for biomedical applications (Roach et al. 2010; Singh et al. 2012).

Another application is super-hydrophobicity (Martines et al. 2005; Bruzzone et al. 2008; Malshe et al. 2013). Super-hydrophobic surfaces can consist of patterns of pillars or grids which alter the contact angle of any liquid on the surface. Water on the surface will then have a larger contact angle than on traditional surfaces, will bead-up and easily roll off the surface. By altering the size and shape of these features, the contact angle can also be decreased and super-wetting achieved (Extrand et al. 2007). Similarly, Weckenmann and Hartmann 2013 have used structured surfaces consisting of repeated grooves to control ink transfer from anilox rolls.

There are many other applications for structured surfaces including: control of boiling (Bruzzone et al. 2008), producing optical metamaterials (Pendry et al. 2004; Leskova et al. 2007) and as abrasive surfaces (Evans and Bryan 1999). In all cases the pattern and shape of features must be controlled to achieve the desired function. Therefore, measurement of the surface is required to validate that the surface has the correct shape.

2.3 Modelling of structured surfaces

On advantage of structured surfaces over stochastic surfaces is that it is often easier to model the functional effect of deterministic structures. While this is not always the case easy modelling is advantageous as it allows it allows a wider range of parameters to be considered than could easily be manufactured. If these models can be validated using a few measurements it is possible to optimise the surface relatively easily. Many examples of modelling of structured surfaces exist. Etsion et al. (Etsion and Burstein 1996; Shinkarenko et al. 2009; Etsion 2013) have developed models for the reduction

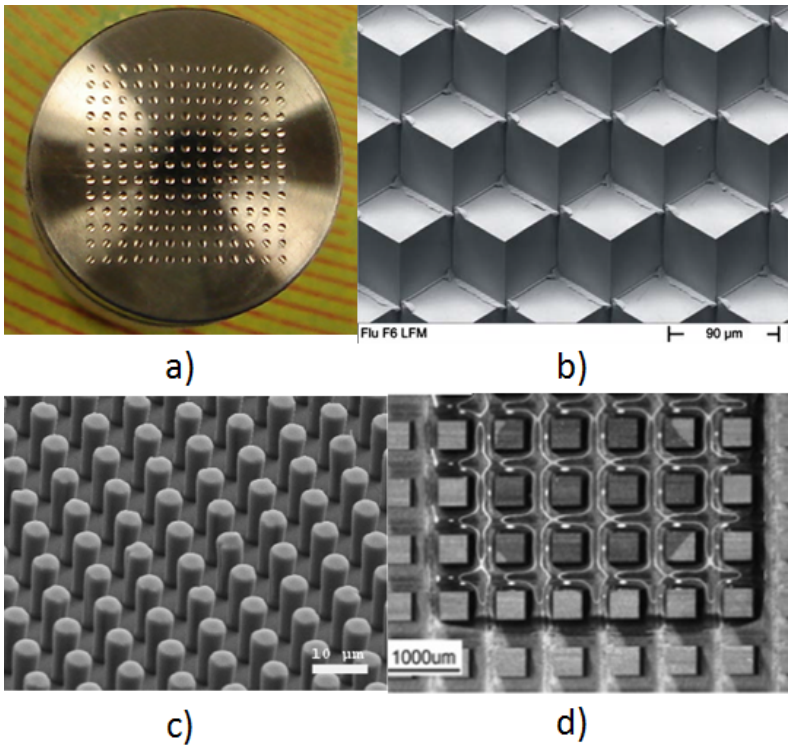

Figure 2.4: Example structured surfaces for optical and hydrophobic effects: a) Micro-lens array (Kong et al. 2010), b) Retro-reflector mold (Brinksmeier et al. 2008), c) A hydrophobic surface manufactured by photolithography (Greiner et al. 2007), d) A wetted hydrophilic surface (Extrand et al. 2007).

.

in friction caused by regular micro-pits in a sliding surface. Similarly optical modelling of structured components is relatively common, Huang et al. (Huang et al. 2007) used optical modelling to support their measurements of a anti-reflectance surface.

The limitations of modelling should be considered. Numerical models, necessarily, simplify reality the significance of these simplifications and the accuracy will inevitably vary from case to case. However, they cannot be discounted and some support from real measurements is needed to validate the results. I such cases accurate measurement of the surface is needed to help validate the model.

2.4 Introduction to surface metrology

Metrology is the science of measurement. For structured surfaces, there are two main reasons to be interested in measurement. Firstly, the determination of functional relationships, which is generally a relationship between the geometric properties of the surface, such as feature size and shape and the functional performance of the surface: friction, hydrophobicity, etc. In most cases, in order to develop such a relationship, it is necessary to be able to measure both the geometric properties and functional performance (although functional performance is not dealt with in this thesis). Once a relationship is determined, an optimal design can be manufactured. As no manufacturing process can be entirely accurate, the design should have some tolerances within which it can be made and still perform acceptably. To determine whether manufactured components fall within these tolerances, measurement is again needed. Without dimensional measurements, performance can only be related to manufacturing conditions, which can cause problems if the manufacturing process changes. Similarly, an unsuitable choice of tolerances, can increase manufacturing cost, for too tight a tolerance, or reduce performance, for loose tolerances. Surface metrology focuses on measurements of surfaces and the deviations from their intended shape (Whitehouse 2010). In particular, the focus is normally on assessment of surface texture, which excludes the very long wavelength form components (Whitehouse 2010). Due to the micro or nano-scale of individual features and large lateral extent of features over a structured surfaces, it is natural that the features of a structured surface are considered part of the surface texture.

2.4.1 Instrumentation for surface metrology

While not the focus of this thesis, some understanding of the instrumentation used in surface metrology is useful. Full details of how various surface texture instruments work can be found elsewhere, for example Leach 2011; Stout and Blunt 2000; Whitehouse 2010. Section 3.2 gives further details on the set-up and operating principles of the instruments used in this thesis.

Surface metrology instruments can be broadly divided into profile and areal instruments. Profile instruments measure a single row of points, i.e. $z(x)$, while areal instruments measure many rows or an image, $z(x, y)$. For structured surfaces, areal instruments are of most interest as the shape of most structures depends on both x and y position. Figure 2.5 shows a possible schema for categorising areal instruments, although other schema are possible. At the highest level, instruments can be divided into scanning probe

and optical techniques.

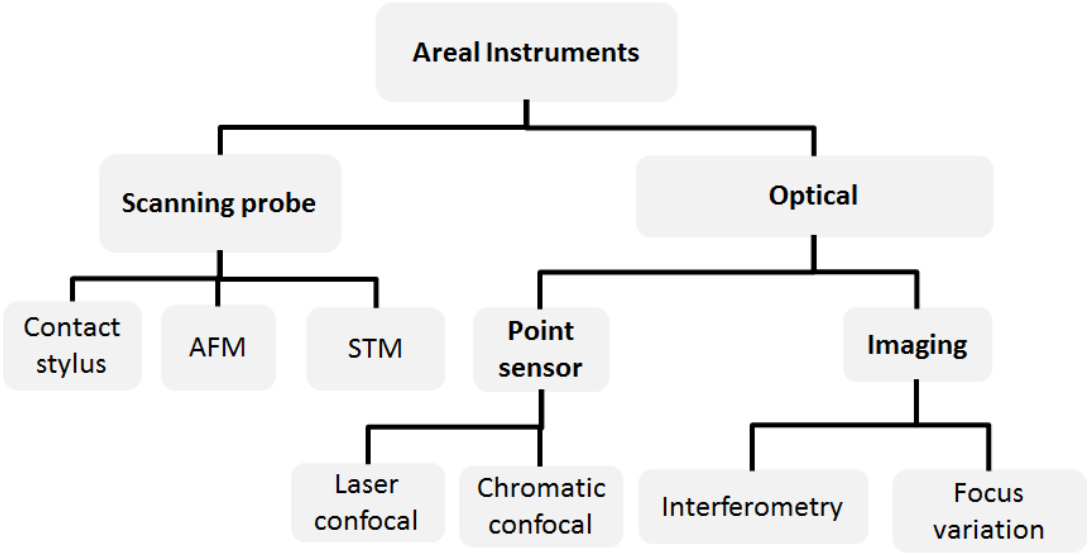

Figure 2.5: Schema to categorise areal surface topography instruments

Scanning probe techniques move a physical probe over the surface to determine its height. Several varieties of scanning probe instruments exist of which the contacting stylus is one of the most common (Whitehouse 2010; Whitehouse 1997). This instrument runs a stylus with a small tip, with a radius of normally $2-10$ µm, over the surface and detects the change in height to create a profile (Whitehouse 2010; Stout and Blunt 2000; Leach 2014). By taking multiple profiles with a known spacing between them an areal map of the surface can be created. Stylus instruments are one of the oldest types of surface texture instrument and can be easily modelled by a ball moving across the surface. However, the stylus must move relatively slowly across the surface and can cause surface damage if the load is not carefully controlled. Other types of scanning probe instruments are the atomic force microscope (AFM) and scanning tunnelling microscope (STM) (Mainsah et al. 2001; Stout and Blunt 2000; Leach 2014). These instruments are designed to measure very small features. However, their bandwidth is limited and they cannot measure larger scale structures. Therefore they are of little interest to measure the surfaces used in this thesis.

Another broad category of instruments is optical instruments. They can be divided into imaging instruments, which use a CCD or similar to measure the entire field of view in a similar way to a camera, and point measuring instruments, which only measure a single point at a time and scan across the surface in a similar way to scanning probe

instruments.

There are various imaging techniques that can be used, of which interferometric techniques are one of the most common. Interferometry relies on interference between light travelling down different paths, which creates alternating patterns of constructive and destructive interference as the different paths move in and out of phase. A classical interferometer uses monochromatic light, resulting in a sine wave intensity pattern as the path length varies. Such an approach is of limited use for measuring rough surfaces or large steps, as many different heights can give the same response. Some method to deal with this ambiguity is needed. For rough surfaces, the most widely used technique is the coherence scanning (white light) interferometer (Groot 2011; Stout and Blunt 2000), which uses a white light source rather than a monochromatic one. The short coherence length of the white light means that constructive interference only occurs when the path lengths are the same. Therefore, by scanning the object vertically, the height of each pixel can be determined by searching for the interference peak. Interferometry is popular as it can give very high vertical resolutions. However, the maximum slope that can be measured is limited by the numerical aperture (NA) of the lens, which for interferometers the NA is often limited to lower NA than other optical instruments, due to difficulties in manufacturing high NA interferometric lenses. Another commonly used imaging technique is the focus variation microscope (Danzl et al. 2011; Helmli 2011), which uses an optical microscope to determine surface height. By scanning in z, the height of the surface can be determined by calculating at which height each pixel is most in focus. This approach requires enough contrast between pixels to determine the focus, which generally limits the application to rough surfaces. However, it only needs diffuse reflection and can, therefore, measure high slopes beyond the NA of the lens.

An alternative approach is point scanning instruments, which only determine the height at a single point and must be scanned across the surface in a similar way to a stylus instrument. One commonly used instrument of this type is the laser confocal microscope (Artigas 2011), which relies on the confocal effect, whereby the laser is reflected off the surface and focused on a pinhole. When the light focused on the surface it passes through the pinhole, producing a high signal. However, if the light is not properly focused on the surface, it will not properly focus on the pinhole and will be blocked producing a low signal. By scanning in z, the height of the point can be determined by finding the peak intensity. Laser confocal instruments generally have slightly lower vertical resolution than equivalent interferometers, as the confocal only has intensity information, giving a wider peak. However, they can also use higher NA lens allowing for measurements of greater

slopes. A related approach is the chromatic confocal microscope (Blateyron 2011). If a white light source is used instead of a laser, a lens with defined chromatic aberration can be used, so that the different wavelengths of light will have different focal lengths. As the distance to the surface changes, the pinhole will block all except a specific colour of light, determined by the height of the surface, which can be measured using a spectrometer. This method can be very fast as it does not need to scan in z. However, the range that can be measured is limited by the optical set-up used (Blateyron 2011).

As well as the instruments described here, there are many others which could be used for surface measurement. In particular, area integrating methods, such as capacitance measurements and scatterometry (Vorburger et al. 2011; Leach 2014; Whitehouse 2010), are not considered as they do not provide a direct measurement of surface topography, but a value that depends on the average topography (and other factors) over an area. In this thesis, the measurement result is always viewed as a height map consisting of an array of surface heights at nominally regular spacing. While in reality the spacing between points may not be entirely regular, either due to variations in pixel spacing across the CCD, uncertainty in the distance moved by the probe between measurements or some other effect, it greatly simplifies calculations to view the pixels as a regular grid as the data can then be treated as a 2D array.

2.4.2 Preprocessing of areal data

Before measured data can be analysed, it must be preprocessed to treat incorrectly measured points and put the data in a more useful form for analysis. The exact steps involved in preprocessing will vary depending on the instrument and surface being measured but the general principles are outlined below.

Treatment of non-measured points

Areal topography measuring instruments often fail to measure the entire surface. Optical instruments, in particular, may fail to collect enough information to determine the height of a pixel, for example due to low reflectance or high slopes, in which case the instrument may mark that pixel as non-measured (Senin and Blunt 2013). Some consistent method is needed to treat non-measured points so that the data can be used. There are two general strategies that can be adopted to deal with non-measured points. Either they can be left as non-measured, in which case algorithms capable of excluding them are needed, or they can be replaced with a plausible height value, computed from interpolation of neighbouring

points. The first approach would be preferable from a metrological standpoint, because it would not introduce additional error in the data. However, adapting data analysis and processing algorithms so that they can discriminate between measured and non-measured points (e.g. through masking solutions) is generally not straightforward and performance may suffer (Senin and Blunt 2013; Senin et al. 2012a). Therefore, correction of non-measured points is often a simpler solution. There are many standard approaches that can be used, usually based on interpolation or fitting of the surrounding valid pixels to determine a best fit value. Notable examples include: median interpolation (Gonzalez and Woods 2008; Senin et al. 2012b), splines (Seewig 2013) and kriging (Raid et al. 2013).

Treatment of measurement artefacts

As well as non-measured points, areal topography measuring instruments may incorrectly determine the surface height for some pixels. A typical example is the bat-wing formations generated by coherence scanning interferometers when encountering a step which is less than the coherence length of the source (Gao et al. 2008; Harasaki and Wyant 2000). These errors, known as measurement artefacts, often look like false topographic formations (e.g. peaks) which are characterised by height values that significantly differ from their immediate surroundings, and can be recognized by outlier-detection techniques.

Unlike non-measured points, measurement artefacts are not easily differentiated from the rest of the measurement. Therefore, an algorithmic approach is required to identify measurement artefacts before treating them. For situations where it is possible to identify measurement artefacts through the detection of outliers, multiple outlier detection techniques derived from statistics can be applied. The standard approach consists of estimating the probability density function of the measured pixels and then identifying those pixels whose heights can be classified as outliers with respect to such a distribution. An example of this approach is Grubbs' test (Grubbs 1950) which marks the pixel furthest from the mean as an outlier until the remaining pixels agree with a students t-distribution with $N-2$ degrees of freedom (N is the number of remaining pixels in the image). Many other similar algorithms exist, for example Le Goic et al. 2013; Ismail et al. 2010; Hodge and Austin 2004.

One issue with such approaches is that they rely on the underlying assumptions of the outlier identification techniques. For flat surfaces consisting of uni-modal, approximately Gaussian local height distributions, identifying outliers is generally simple; the ideal case being a horizontal rough surface with no significant high-scale formations. On the other

hand, even the simplest step, or the presence of an underlying non-flat form or long wavelength components, complicates the situation and makes the detection of outliers more difficult. Unfortunately, structured surfaces often contain such components. The typical solution to this problem is to remove form and long wavelength components from the topography, by subtracting a filtered version of the surface from the original. The resulting surface will contain only the higher frequencies. The filtered topography can be obtained in a variety of ways, such as polynomial fitting or by using Gaussian or median filter (Senin et al. 2012b; Gonzalez and Woods 2008). However, polynomial fitting does not work well in the presence of steps and other sharp discontinuities. Another recent approach by Le Goic et al. 2013 is to use discrete modal decomposition, which decomposes the surface into different modes, in a similar way to the Fourier or wavelet transform. Once the residual, high frequency surface is obtained, outliers can be searched for. Once identified, outlier pixels can either be excluded or corrected to more plausible height values, as for non-measured points.

Filtering

It is often necessary to filter a measured surface to reduce high frequency noise, which is outside the measurement bandwidth of the instrument being used. For stochastic surfaces a Gaussian filter is commonly used, which has the advantage of being a linear filter, so the frequency response is constant. Additionally, the Gaussian filter has a monotonic frequency response and does not produce ringing effects that many other linear filters, such as the mean filter do (Seewig 2013; Brinkmann et al. 2001). Figure 2.6 shows an example of the Gaussian filter and its frequency response. The shape of a Gaussian kernel can be given by

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2.1}$$

where $g(x, y)$ is the amplitude at point (x, y) from the centre of the filter and σ is the standard deviation of the filter. The standard deviation of the filter determines its spatial (xy) extent and, therefore, the strength of the filter. However, in surface metrology, it is more common to consider the size of a filter in terms of a cut-off wavelength, λ_c where the frequency response of the filter drops below 50 %. The frequency response can be thought of as the amplitude of a sine wave of given frequency or wavelength that is measured after filtering compared to the input. Alternatively, the frequency response can be assessed in terms of the Fourier transform of the filter. The Gaussian filter can also be written in

terms of the cut-off frequency such that (Seewig 2013)

$$g(x,y) = \frac{\pi}{\ln(2)\lambda_c^2} e^{-\frac{\pi^2}{\ln(2)\lambda_c^2}\left(x^2+y^2\right)} \tag{2.2}$$

In which case σ and λ_c are closely related with

$$\sigma = \sqrt{\frac{\ln(2)}{2}}\frac{\lambda_c}{\pi} \tag{2.3}$$

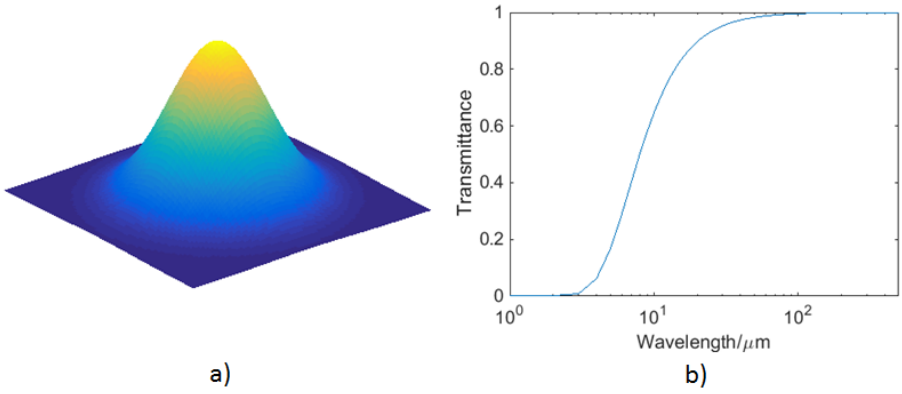

Figure 2.6: a) Kernel of 2D Gaussian filter and b) frequency response for 8 µm *cut-off Gaussian filter*

The standard Gaussian filter, as described above, will have issues at the edges of the measured area. Towards the edge of the image, some points in the filter kernel will lie outside the measured area. If this is the case, the filter will not return a valid result for these points. Padding using either a constant value or replicated data is often used to allow the filtered data to be calculated near the edge of the image. However, this approach can still skew filtering, producing distorted results. Another, more effective, method of correcting for edge effects is to modify the filter kernel where it would otherwise cross the edge of the image, such that the total amplitude of the kernel is maintained (Seewig 2013), see figure 2.7.

For structured surfaces, Gaussian filters may not be ideally suited, as they can distort the position of edges in the image and can therefore distort the features of the surface (Bergholm 1987; Zhong and Ma 2010). If the features are large compared to the size of the smoothing kernel used, then this distortion will be relatively small. However, as the feature size decreases, the distortion effect will become more and more significant. Therefore, other, edge preserving, filtering techniques are considered.

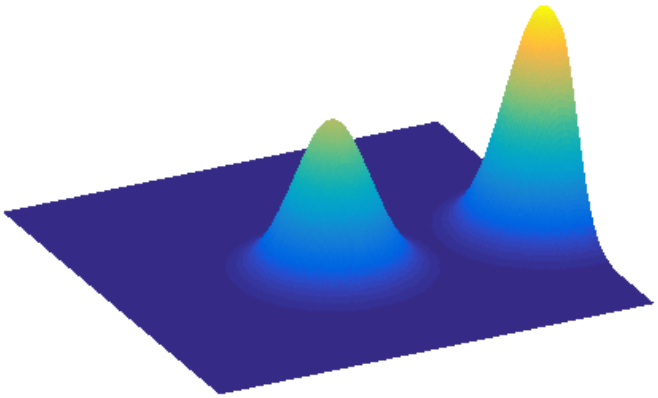

Figure 2.7: Example of Gaussian kernel and modified kernel as it crosses the image boundary given by the edge corrected Gaussian filter.

One edge preserving filter is the morphological filter (Gonzalez and Woods 2008; Seewig 2013). This filter is based on applying morphological opening and closing operators to the surface, with spherical or flat structuring element. While it is superior to the Gaussian filter in terms of preserving significant edges, the conventional algorithms are computationally intensive, especially for large structuring elements. This issue can make the filter prohibitively slow to apply, especially for large images. Recent research (Lou et al. 2013) has investigated novel algorithms to perform morphological filtration which show promise to achieve much faster computation.

Another possible filter, proposed for use on structured surfaces by Jiang and Whitehouse 2012, is the anisotropic diffusion filter, which has been used for filtering objects in image processing applications (Perona and Malik 1990; Gerig et al. 1992; Bovik 2010). This filter results in efficient smoothing of low-gradient regions while at the same time, sharp edges are preserved.

Levelling and form removal

Many surfaces have a large scale form that is approximately planar. It is important to level these surfaces such that the surface is flat and the mean is at zero, as these assumptions are used in calculation of most parameters. One way to do this for surfaces

with small tilts is by subtraction of a least squares mean plane. The least squares plane of the unlevelled surface is calculated, then for each pixel the corresponding value on the plane is subtracted, such that the least squares plane of the resulting surface is at $z = 0$. The disadvantage of this technique is that it will distort the surface, see figure 2.8. A more robust approach is to calculate the slope of the fitted plane and perform a rotation on the surface such that the plane is flat. In which case, no distortion in the surface occurs (Senin and Blunt 2013). However, rotation is far more computationally intensive than subtraction and will change the pixel spacing, which can cause complications in later analysis. Therefore, for small angles, where the distortion caused is small, subtraction is preferred and the distortion is ignored.

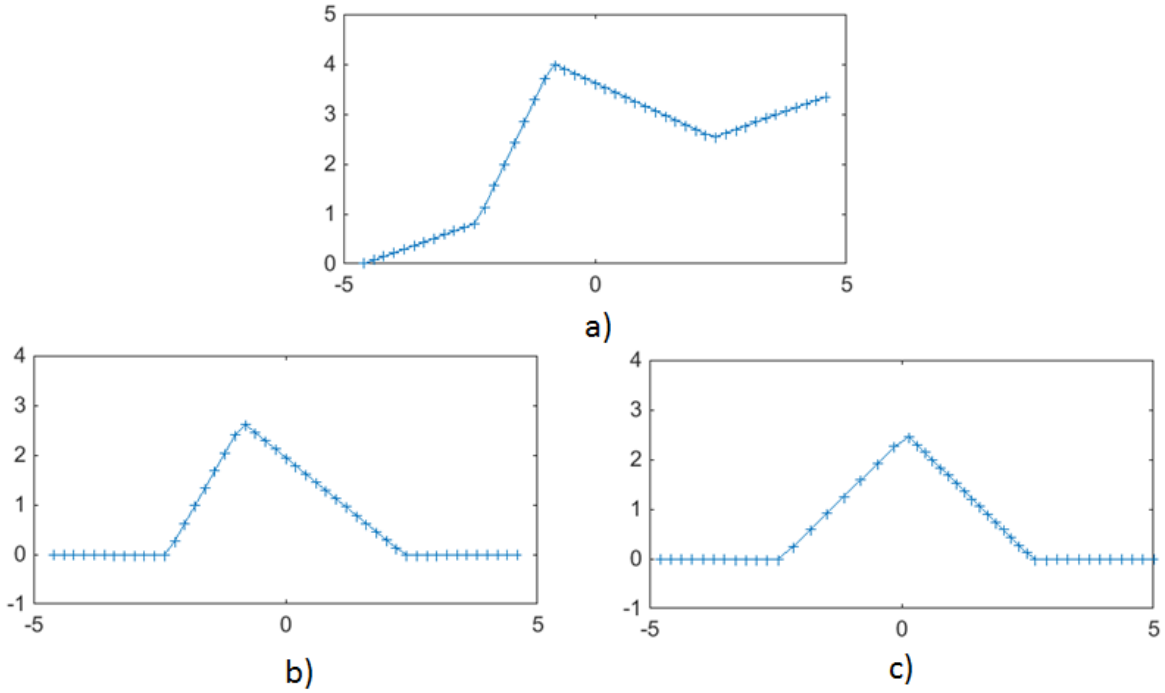

Figure 2.8: Example of errors caused by levelling with subtraction vs rotation: a) Simulated profile, of triangular feature tilted at 20°. b) Profile levelled by subtraction, feature is distorted but constant spacing is maintained. c) Profile levelled by rotation, feature is not distorted but pixel spacing is not constant.

The standard levelling approach described above was developed for use with stochastic surfaces and assumes that small-scale roughness is the only deviation from the plane. However, for structured surfaces this is not the case. When levelling structured surfaces, it is necessary to exclude the surface features from the least-squares fitting, as otherwise

they can skew the mean plane (Senin and Blunt 2013), as shown in figure 2.9. It is important to note that the plane should still be subtracted from the entire surface, to ensure a single reference plane for the entire surface. Fitting a least-squares plane to the background surface implies that the features have been identified and their position is known, for example using one of the segmentation methods described in section 2.6. However, the segmentation techniques often rely on having well levelled images. Therefore, an iterative approach is recommended whereby the image is segmented, then levelled, then segmented again until a stable image is reached (Senin et al. 2014). An additional issue is that selectively fitting to the background is not possible for structured surfaces where the features contain the entirety of the tiles, such as in figure 2.2(b). Levelling against the whole surface may still be unsatisfactory if all the surface pixels are used, as partial tiles are likely to be present at the image borders and these could skew the least-squares mean plane.

A similar approach could be taken when dealing with a non-planar form. The general approach is to separate the form from the surface texture when calculating parameters. Therefore, rather than subtracting a plane from the surface, a higher order polynomial can be subtracted.

2.4.3 Areal surface texture parameters

Traditionally, the focus of surface metrology has been on the characterisation of unstructured, or stochastic, surfaces. For stochastic surfaces it is of little use to characterise the geometrical size and shape of individual features as would be done for larger scale dimensional metrology (Whitehouse 2010). Instead, parameters are defined based on statistical properties of the topography and these parameters are used to characterise the surface. For areal data, a standard set of parameters are specified in ISO 25178-2 (ISO 2012). This standard, which builds on previous standards for profile data (ISO 1997), specifies a number of parameters in two broad categories: field and feature parameters. Table 2.1 provides a description of some of these parameters.

Field parameters are the most commonly used category of parameter and use all the data on the surface without distinction between points. ISO 25178-2 breaks field parameters into four sub-groups: height, spatial, hybrid and functional parameters. For a full description of field parameters see ISO 25178-2 (ISO 2012) or Blateyron 2013b. Height parameters consider only the heights of surface points, not their position, and consist of statistical moments of the surface as well as extreme points. For example, Sq

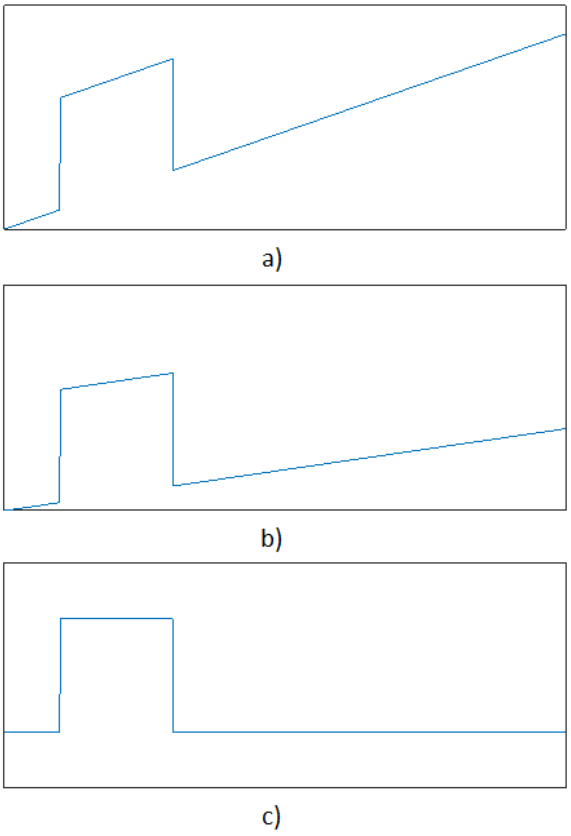

Figure 2.9: Effect of excluding features when levelling. a) An example profile containing a feature measured at a slope. b) Profile levelled using all points. As feature is off-centre it skews the levelling and resulting surface is not well levelled. c) Profile levelled excluding the feature from the fit, resulting in a well levelled surface.

is the standard deviation of the surface (the second moment), Ssk is the skewness (third moment), Sku is the kurtosis (fourth moment) and Sz is the difference between the maximum and minimum points on the surface. Spatial parameters are concerned with the autocorrelation function and two are given: the autocorrelation length, Sal, and the texture aspect ratio, Str. Similarly there are two hybrid parameters defined: Sdq and Sdr; these are based on the average gradient over the surface. The functional parameters are based on calculating the material ratio curve of the surface (Blateyron 2013b). A range of parameters are then defined based on properties of this curve.

Feature parameters are an alternative set of parameters, developed from the idea of motifs (Dietzsch et al. 1998; Stout and Blunt 2000) used for profile measurements. Rather than consider the entire surface indiscriminately, the surface is divided into different

Table 2.1: Descriptions and formula for some surface texture parameters.

Parameter	Description	Formula and Notes
Sa	Arithmetic mean height	$\frac{1}{A}\sum_A \lvert z(x,y)\rvert$
Sq	Root mean square height	$\sqrt{\frac{1}{A}\sum_A z(x,y)^2}$
Ssk	Skewness of the surface	$\frac{1}{Sq^3}\frac{1}{A}\sum_A z(x,y)^3$
Sku	Kurtosis of the surface	$\frac{1}{Sq^4}\frac{1}{A}\sum_A z(x,y)^4$
Sp	Maximum peak height	
Sv	Maximum valley depth	
Sz	Maximum height of the surface	$Sp + Sv$
Sal	Autocorrelation length	Horizontal distance of ACF which has fastest decay to value, s
$Smr(c)$	Material ratio of the surface	Percentage of surface below height, c
Sk	Core surface height	Calculate using material ratio curve
Svk	Reduced valley depth	Calculate using material ratio curve
Spd	Density of peaks	Calculate on segmented surface
$S10z$	Ten point height of surface	Sum of average of five highest and lowest points after segmentation

regions or features. Properties of these features can be calculated and feature parameters are the statistics of these properties. While feature parameters should be considered more a toolbox than a fixed set of parameters (ISO 2012), some commonly used parameters are named for example, Spd is the mean peak density and $Sp5$ is the mean of the five highest peaks. When calculating feature parameters the biggest concern is how to determine the features. ISO 25178-2 suggests using morphological segmentation and pruning to segment the surface. This approach is discussed in more detail in section 2.6.2, or see Blateyron 2013a.

2.5 Approaches to characterisation of structured surfaces

Structured surfaces are formed of micro or nano-scale tiles repeated many times, so that the resulting pattern covers an area orders of magnitude larger than the scales of the individual features. Function in structured surfaces is generally exploited by the combined interaction of numerous tiles (Evans and Bryan 1999; Bruzzone et al. 2008). Therefore, the final goal of a procedure for characterising a structured surface is the assessment of properties pertaining to the pattern as a whole, rather than the characterisation of a single tile or feature. Conceptually, this is consistent with the approach adopted in surface metrology for characterising conventional surface texture, where the ultimate goal is the identification of surface texture parameters, which describe information pertaining to the whole topography (ISO 2012). However, although the final goal of the characterisation is the same, the approaches adopted in the literature to obtain such results for structured surfaces have been varied, sometimes departing significantly from those adopted for more conventional stochastic surfaces. While a detailed illustration of the main existing approaches will be provided in the following sections, a preliminary taxonomy is discussed below and shown in figure 2.10. Regardless of the characterisation method, it is assumed that surface topography data has been obtained by an areal topography measuring instrument and is available as a height map, consisting of a set of pixels arranged into a regular grid.

Two general categories of characterisation methods can be considered. The first group of methods views structured surfaces in a similar manner to how stochastic surfaces would normally be assessed, i.e. field parameters are calculated based on statistical properties of the entire measured area. These parameters can either be based on the standard areal surface texture field parameters, described in ISO 25178-2 (ISO 2012), or on other, custom parameters. These methods consider all measured pixels similarly, with no differentiation between pixels and, therefore, ignore the tile structure of structured surfaces.

The second group encompasses many different methods, sharing a common trait in that they introduce differentiation between different regions of the surface. In these methods, a few key steps are always recognisable: the surface is segmented so that the individual tiles and features are isolated, and parameters are computed starting from the tiles and features. Once many individual features have been classified parameters can trivially be constructed based on the statistics of the individual features.

There are a variety of ways of performing segmentation and feature identification, and

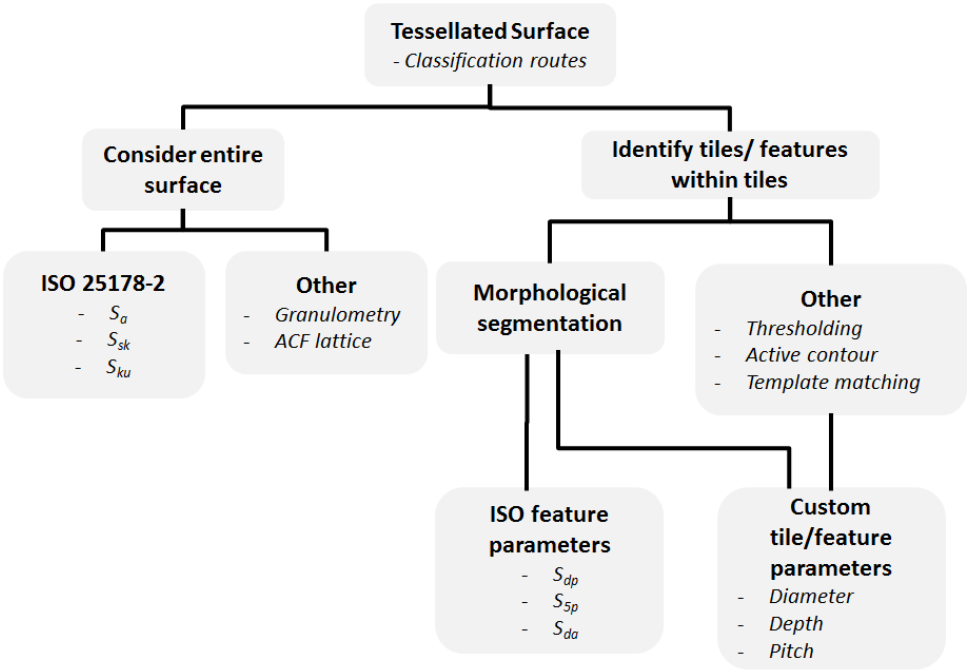

Figure 2.10: Overview of the classification taxonomy for structured surfaces

there are a variety of ways to encode the results through parameters. One of the most common methods consists of using morphological segmentation, and then computing the related ISO feature parameters (ISO 2012). However, many different, non-ISO methods have been explored to segment the surface into tiles and features, and to compute parameters that describe the segmented surface. Some of these methods target the tiles, others the features within the tiles. In either case, the parameters proposed in the literature generally try to capture either tile and feature-related properties (i.e. tile and feature shape and size), or lattice-related properties (i.e. distance between and alignment of the tiles and features). These two different approaches to characterisation are now discussed in more detail.

2.5.1 Characterisation without segmentation

This section discusses characterisation methods aimed at structured surfaces, which do not require the topography to be segmented in order to identify tiles and the functionally relevant features located within them. With no segmentation, pixels are all treated equally, and included in the computation of the parameters. The techniques belonging to this category can be divided into those adopting ISO parameters for describing the topography and those that are based on devising custom parameters instead.

ISO field parameters

Some researchers have attempted to characterise structured surfaces using ISO field parameters. In particular, Podgornik and Sedlacek 2012 have tried, with some degree of success, to find a correlation between the kurtosis and skewness of laser textured dimples and grooves, and their coefficient of friction under boundary and mixed lubrication conditions. They succeeded in showing correlation between skewness, kurtosis parameters and the friction coefficient. Such research results demonstrate that standard surface texture parameters may be useful for characterisation of structured surfaces. However, it must be questioned whether similar relationships between field parameters and function could be derived for similar unstructured surfaces. Indeed, Podgornik and Sedaleck note that the relationship they obtained follows the same pattern as that observed for friction in conventionally machined surfaces. Therefore, while a correlation was found between the surface texture parameters and friction, no information could be obtained about how structuring affects the functional performance. This is because many different feature shapes and layouts may lead to the same height distribution and, therefore, to the same field parameters.

Accordingly, while ISO field parameters are widely used in analysis of surface texture for stochastic surfaces, several researchers have questioned their suitability for the analysis of structured surfaces. Weckenmann and Hartmann 2013 claim that as profile and areal field parameters assume the surface has a random topography, such parameters are poorly suited to the analysis of surfaces with micro and nano-scale structures and fail to map the geometry of the micro-structures successfully onto the functional performance of the surface. Therefore, many micro-structured components can currently only be verified by functional tests. Similarly, Blunt and Xiao 2011 argue that conventional surface texture parameters are designed to determine statistical properties of the entire surface, whereas for structured surfaces the primary interest is in the deviations from the nominal of the individual features.

ISO field parameters of a structured surface may, in certain cases, partially succeed in correlating with functional performance or the size of individual feature. However, such relationships do not uniquely define the surface and can be affected by many other factors.

Custom parameters

Researchers have also attempted to define custom parameters in order to obtain a stronger relationship between the parameter and functional performance, or even with the actual geometric properties of the features forming the pattern. Of particular note, Geringer et al. 2013 propose a method based on a 3D analogy of granulometry. By calculating the volume of the surface after morphological opening and closing operations with a range of structuring element sizes, a relationship between volume and structuring element size can be created. In this relationship, the structuring element size corresponding to the maximum rate of change of the volume curve was found to be related to the characteristic lateral size of the dominant features in the pattern. Geringer et al. tested this approach on a laser textured surface and showed a good correlation between the calculated and actual lateral feature size.

While Geringer's method appears promising for the characterisation of the specific features tested, care must be taken when applying this method to other surfaces. All the surfaces used by Geringer consist of features that are approximately circular in shape. The behaviour when other surfaces are considered has not been investigated. In particular, it appears that if severely elongated features were used, then only the smallest dimension of the feature would be detected correctly. Additionally, this approach is limited to considering the lateral size of the feature. While deeper features will give a stronger response, due to a greater change in volume, there is no way of directly determining feature depth or other properties which may be of interest.

Another approach, suggested by Zeng et al. 2013, uses the autocorrelation function (ACF) to find the lattice properties of a structured surface by finding the translation vectors between peaks in the ACF. The translation vectors can be used to define parameters describing layout properties. Due to the regular nature of structured surfaces, the ACF should consist of a number of regularly shaped, sharp peaks. By measuring the distance between such peaks, the average lattice properties of the surface can be calculated.

While the custom parameters discussed above show some promise in characterising structured surfaces, they have similar drawbacks to using ISO field parameters in that they have limited direct relationship to the geometry of the features. Therefore, these parameters are easily affected by factors other than those they aim to characterise. This limitation makes it difficult to generalise such approaches to different surfaces.

2.5.2 Characterisation based on segmenting the topography

Many characterisation methods for structured surfaces are based on identifying the features and tiles of the surface. This identification implies that the surface is segmented into different regions. While an overview of the approaches and parameters used are discussed in this section, detailed consideration of segmentation methods is given in section 2.6. Once the segmentation has been performed, the final characterisation of the surface is based on computing parameters that capture the properties pertaining to the isolated tiles and features and their spatial layout.

ISO morphological segmentation and feature parameters

Morphological segmentation is currently the only segmentation method recommended by ISO 25178-2 (ISO 2012). In morphological segmentation the topography is segmented into topologically connected hills or dales (Senin and Blunt 2013; Gonzalez and Woods 2008; Xiao et al. 2006). After segmentation has been performed, parameters can be computed that capture the properties of such hills or dales; these are known as ISO feature parameters (ISO 2012). For structured surfaces, the ISO approach to segmenting and computing parameters may be applicable, as long as a hill or dale segmentation can be appropriately configured to capture the tiles and the features contained within. This is a non-trivial problem, as illustrated by Senin et al. 2013. In addition, even when each tile or feature is appropriately modelled by a subset of the segmentation regions, it is still necessary to identify those feature parameters that better capture the properties of the tiles and of their layout. If successful, the combination of ISO morphological segmentation and ISO feature parameters allows the topography to be described in standardised and repeatable terms.

Hartmann and Loderer 2013 have used morphological segmentation to identify the features in dimpled surfaces for friction reduction. Once features are identified their diameters and depths are calculated to produce distributions of diameter and depth for the entire surface.

Blunt and Xiao 2011 and Blunt and Scott 2012 have considered the characterisation of laser textured hard disk drives with a tessellating pattern of bumps. The Laplacian of a Gaussian (LoG) operator is used before ISO morphological segmentation on the resulting image, to ensure that the segmented regions are representative of the real bumps. Parameters, such as average diameter and depth of the bumps and average separation between bumps, are then assessed.

An alternative strategy, still involving an ISO-compliant segmentation process, consists of applying a segmentation technique not directly to the original topography of the structured surface, but to an ancillary surface obtained from it. Jiang et al. 2007 and Jiang and Whitehouse 2012 have applied ISO morphological segmentation to the ACF of a structured surface. This segmentation leads to a robust identification of the general shape of the tile and, therefore, lattice parameters, because dale segmentation correctly identifies the shape of the region comprised within adjacent peaks in the ACF.

While the use of morphological segmentation and ISO feature parameters show some promise for characterising structured surfaces, other methods are also possible which may be better suited for some surfaces.

Other segmentation techniques and custom parameters

Many other methods have been used to segment the topography of a structured surface; they are discussed in detail in section 2.6. Accordingly, researchers have developed various custom parameters to describe the pattern units and their spatial layouts as obtained after the segmentation. For example, Kong et al. 2010 used a height-based threshold (see section 2.6.1) to identify the lenses in a micro-lens array. Various parameters were then used to characterise the micro-lens array, including the minimum, maximum and standard deviation of the roundness and lattice deviations in both a spatial and angular sense. Zhu 2012 and Zhu et al. 2009a have applied an active contour (see section 2.6.3) based algorithm to the segmentation of a range of surfaces, including etched silicon patterns, laser textured patterns and computer chip pins, and show that this method can accurately identify the surface features. However, the calculation of relevant parameters from these regions is not directly considered.

While custom parameters based on segmentation can be used to identify surface features and characterise them, further consideration must be given to how the choice of segmentation method effects the result as well as the effect of varying the topography.

2.6 Segmentation methods

Segmenting an image divides it into different regions based on some heuristic. Therefore, segmentation can be used to identify tiles and features on a structured surface and their boundaries. In the following section, various segmentation methods will be examined and the application of such methods to feature and tile extraction in the specific case of

structured surfaces will be discussed.

2.6.1 Thresholding

One of the most straightforward and widely used ways to segment surfaces is to threshold the surface based on some local property of the pixels. Under this approach, a threshold, T, is set; all pixels with value less than T are marked as part of a feature and all pixels greater than T are marked as not belonging to a feature, or vice versa as appropriate for the surface. Many different local surface properties, such as height, gradient or local texture parameters, can be used for thresholding with differing effects, as shown in figure 2.11.

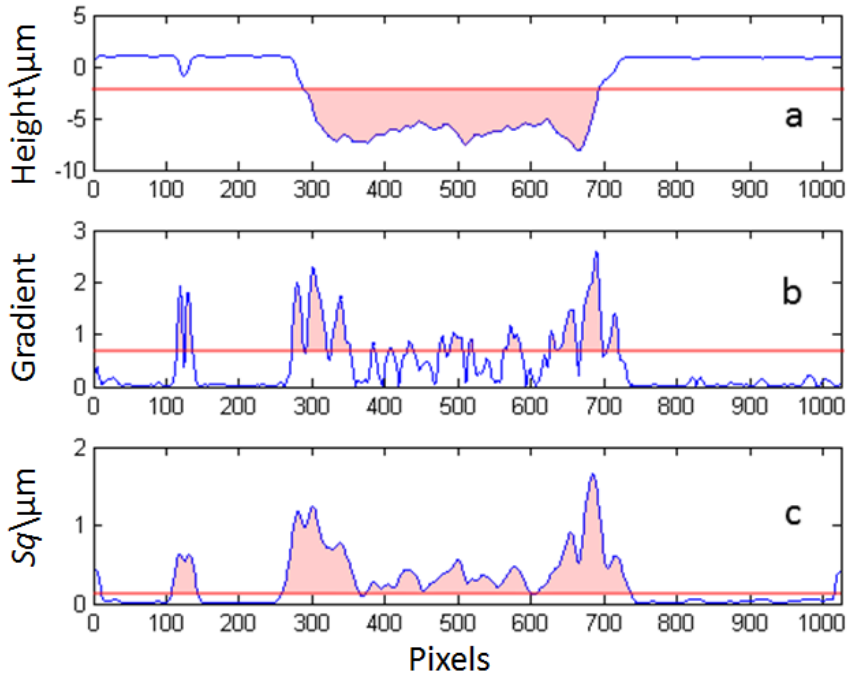

Figure 2.11: Effect of thresholding on example profile using different image properties. The principle is the same for areal data; profiles are shown here for ease of visualisation. In each case the red line shows the threshold value and feature areas are highlighted: a) surface height, b) gradient, c) local Sq in 21×21 *moving window.*

The height of the pixel is probably the most straightforward and widely used threshold. Height thresholds provide an intuitive and computationally simple way to segment the surface. They are most effective when used to detect sharp edges, such as steps, and have

been used extensively in the field of computer vision (Sahoo et al. 1988; Lee et al. 1990). Senin et al. 2012b have recently devised a technique for the identification of thin foil laser targets for ion beam acceleration experiments, which uses a height thresholding operation to discriminate between the target and background surfaces, as shown in figure 2.12. For structured surfaces, Kong et al. 2010 have used a height threshold to identify lenses and separate them from the background in a micro-lens array.

However, there are several issues with height thresholds which limit their applicability. Firstly, it is difficult to select an appropriate threshold value. While several techniques exist to aid in selection of an appropriate threshold, including using the image histogram (Gonzalez and Woods 2008), Otsu's method (Otsu 1975), k-means clustering (Macqueen 1967; Xu and Wunsch 2009; Senin et al. 2007) or using a percentage of a field parameter (Senin and Blunt 2013), these approaches rely on the height distribution being multi-modal and the feature being contained entirely within a well-defined number of modes. If this is not the case, the correct threshold is not well defined (Gonzalez and Woods 2008). Another concern is that height thresholds can produce many spurious features, such as those shown in figure 2.13, either due to noise in the measurement or the roughness of the background surface. While it may be possible to use filtering or other techniques to remove these falsely identified regions, these approaches are unlikely to be perfect and must be tuned for a particular surface. A third serious concern when performing a height threshold is the effect of levelling. Height thresholds can perform poorly when dealing with unlevelled data because slopes can affect how pixels are marked. For example, if the background is tilted, a threshold that successfully detects the feature is likely to select a significant part of the background.

It is also possible to threshold on other local surface properties. One option is to use the value of a surface texture parameter as the threshold. From a metrological point of view, surface texture is generally considered in terms of the areal surface texture parameters described in ISO 25178-2 (ISO 2012) and is calculated as a single value for the entire surface or image. However, if these parameters are calculated in a local region around each pixel, rather than for the whole image, then a value can be attached to each pixel and thresholds can be set to segment the surface. Senin et al. 2007 have used this method and successfully segmented different textured squares and micro-indentations by using k-means clustering to determine thresholds on the local Sq value.

Depending on how texture parameters are calculated, texture-based thresholds may be less susceptible to levelling issues than height thresholds. However, there are additional drawbacks to consider. Texture parameters are calculated on each pixel using information

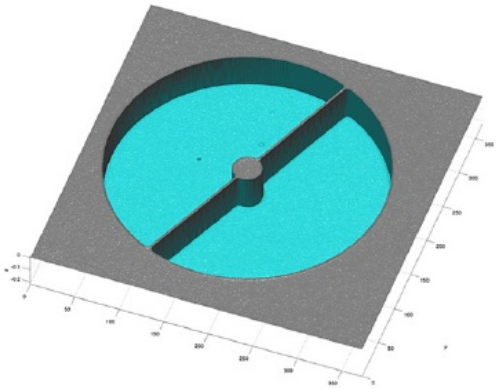

Figure 2.12: Height thresholds can be very effective when the surface contains step-like features (Senin et al. 2012b).

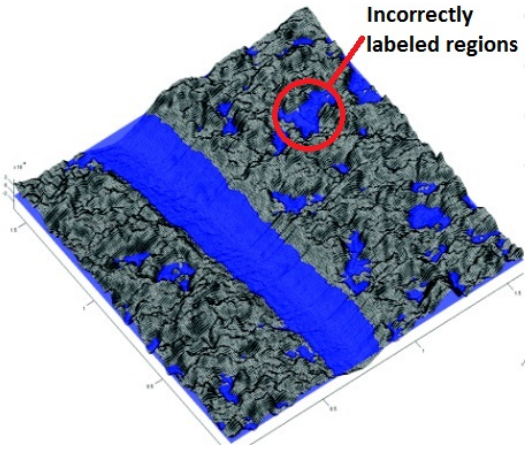

Figure 2.13: When the surface contains shallow features with not sharp edges height thresholding often produces many false features (Senin and Blunt 2013).

from a given amount of neighbouring pixels. Therefore, sharp transitions in the topography, such as a step, will produce different results in terms of the transition of the local texture parameter. This is because at each position the parameter is computed from an aggregation of pixels belonging to both sides of the transition, with varying relative weights as the point moves across the transition itself. The simplest example of this would be the moving average, which results in a smoothed transition when applied to a sharp step. Most texture parameters computed locally over a region of neighbours would have similar behaviour, which must be taken into account.

Additionally, variations of the texture parameter value when computed over boundary pixels must be considered; at the image boundaries, fewer pixels are available to compute the local texture parameters, and this may affect the parameter value. Also, choices of how to obtain the missing points (e.g. symmetric extrapolation, wraparound, constant value padding) affect the parameter, and thus the segmentation results.

Finally, texture thresholds suffer many of the same issues as height thresholds with regard to setting appropriate values. Senin et al. 2007 used k-means clustering, with between two and six clusters to deal with this problem, which was effective for the surfaces they used.

Another common option is to set the threshold based on the gradient of the surface. There are a number of ways to calculate the gradient of a surface. The Sobel filter is a popular method for range images, which convolves the surface with a pair of filter kernels

to produce gradients in the x and y directions (Gonzalez and Woods 2008). These maps can then be combined to produce the gradient magnitude, against which the threshold is applied. Setting such a threshold detects areas of high local slope, which is generally the feature boundary and, therefore, separates the background and feature regions. Thresholding based on the gradient has similar advantages and disadvantages to texture-based thresholds, as they are closely related. In particular, the choice of threshold is still an issue. As before, automated methods can be used. However, one possible advantage of gradient-based thresholds is that it may be slightly easier to set the threshold manually, as in many cases the expected gradient of the background is zero (or at least very small). However, in practice the choice of threshold still has a significant effect on the segmentation boundary. Additionally, the centre of features often have low gradient compared to the boundaries and therefore may not be marked as features, depending on how the threshold is set. As for texture based thresholds, filling of enclosed regions can mitigate this issue.

Threshold based techniques are generally poorly suited to identifying pattern units (tiles) in structured surfaces. As when considering pattern units, all the units will have similar properties and cover the entire surface. Depending on the specific surface, it may be possible to set a threshold such that a region near the boundaries of the texture units is segmented from the rest of the surface, for example, by setting a threshold on the height on the pyramidal surface in figure 2.2(b). However, this is still a non-ideal situation as the resulting boundary region will inevitability not result in a thin region and so high accuracy in determining the tile boundaries will not be possible.

2.6.2 Morphological segmentation

Morphological segmentation is another segmentation approach, which is based on the idea of morphological watersheds, proposed by Maxwell 1870. The idea is similar to the principle of the watershed of a river basin; in that the surface is segmented such that a drop of water placed at any point within a region flows down to the same point. The boundaries between regions are then the watershed lines, where a drop can flow in either direction. Inverting the topography, or equivalently the direction of gravity, leads to morphological segmentation into hills. Figure 2.14 shows a schema of morphological segmentation for profile data. This can easily be generalised to areal data. Several different algorithms for efficient watershed segmentation have been developed (Meyer 1994; Vincent 1993; Bieniek and Moga 2000). These operate either on the basis of immersion simulations (floodfill

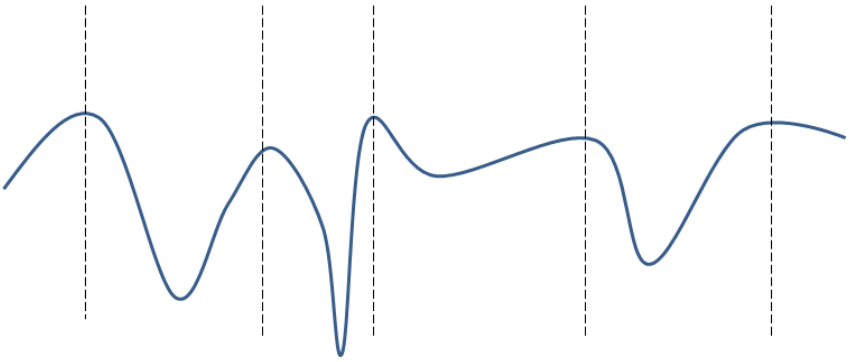

Figure 2.14: Schematic of morphological segmentation on a profile. Watershed lines exist at maxima of the profile (ridge lines for when generalised to areal data).

techniques) (Vincent 1993), where the image is flooded from beneath and watershed lines are defined where two flooded regions meet, or by assessing how each point on the image flows to a minima (rainfall techniques) (Meyer 1994)

One of the major issues with the watershed transform is that it tends to severely over-segment the image, as shown in figure 2.15a on a test surface of a laser-textured micro dimple. This over-segmentation is due to the fact that each local minima in the image will have its own watershed region. Due to noise in the image and roughness of the surface, there will be many such regions which are unlikely to have any relevance to the surface function. Therefore, some method of region merging is required to merge the insignificant regions together into more significant regions, which ideally correspond to the structures on the surface. One of the simplest ways to merge regions is to smooth the image before segmentation using a Gaussian, or some other filter (Gauch 1999; Undeman and Lindeberg 2003), which will reduce the number of minima in the image and, therefore, the number of watershed regions. However, such smoothing will also distort the image and cause the boundaries to move.

Another approach to deal with over-segmentation, often used in conventional image processing, is marker-based segmentation (Gonzalez and Woods 2008; Gao et al. 2004; Sun and He 2008). Marker-based segmentation relies on placing markers on the image which correspond to the significant features. The image is then morphologically deformed such that minima only exist at these markers. This reduces the number of watershed regions and helps them to agree with the significant features. However, marker-based segmentation is not widely used in surface metrology because placement of markers requires significant a priori knowledge about the location of the features, either to place the markers manually or to control some algorithm to place them, which is not gener-

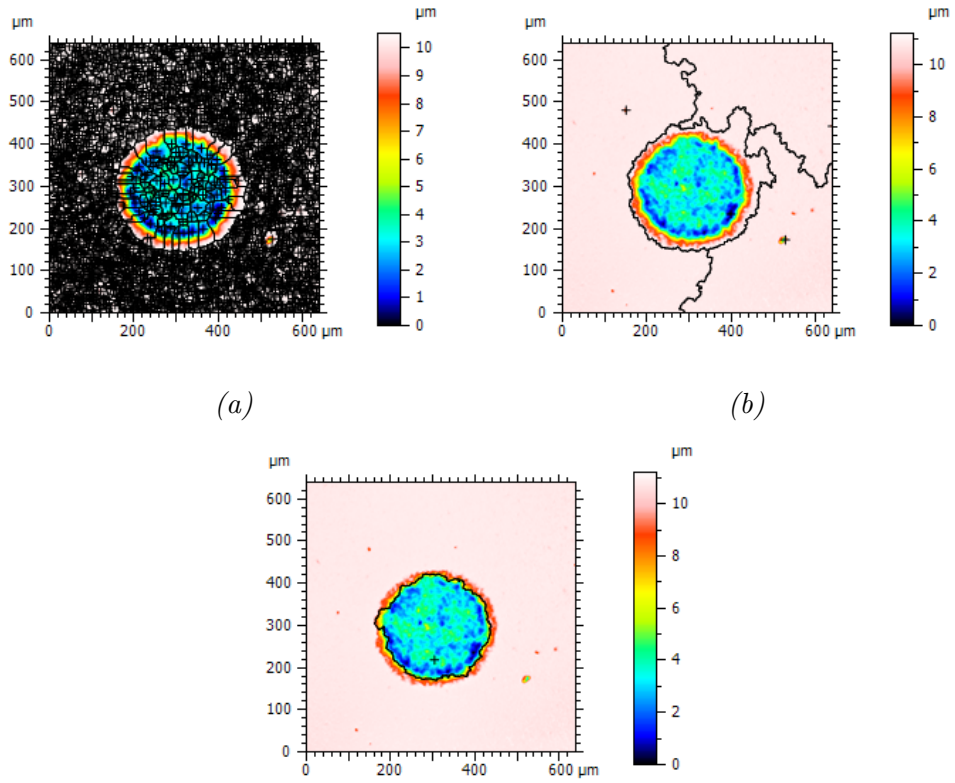

Figure 2.15: Comparison of a laser textured dimple surface, segmented by the watershed transform. a) With no pruning the feature is severely over-segmented. b) After pruning with a height threshold of 20 % of Sz and area threshold of 10 % of the image the feature is segmented but does not match the boundaries well. c) Using the watershed transform on the gradient map and pruning with 5 % of maximum gradient and 10 % area threshold the feature matches the observed boundary far more closely.

ally available for stochastic surfaces. However, for structured surfaces marker placement should be much easier due to their deterministic nature.

A third approach to the over-segmentation problem is region merging techniques, of which several equivalent methods exist (Wolf 1991; Bleau and Leon 2000; Scott 2004). These rely on setting thresholds on properties of the watershed regions, often the depth and area. Watershed regions with values smaller than these parameters are then merged with neighbouring watersheds until all remaining regions are above the threshold. The problem with region merging approaches is that the thresholds must be chosen heuristically, as they will depend on both the desired segmentation and the surface. Therefore,

there is no guarantee that a threshold which was successful for one surface will also be successful for another (Senin et al. 2013).

In surface metrology, morphological segmentation is primarily used as the method to segment a surface before determining feature parameters (ISO 2012; Blateyron 2013a). For this purpose, region merging techniques are used to deal with over-segmentation. ISO 25178-2 (ISO 2012) recommends using Wolf's method (Wolf 1991) for region merging, although it points out that other methods, such as that of Bleau and Leon 2000 are equivalent. Scott 2004 has shown that Wolf's method provides a stable segmentation for determining surface features.

While application of morphological segmentation on height maps has been successful for many surfaces, morphological segmentation can equally be applied to other surface maps, for example the gradient map (Senin et al. 2013; Jackway 1996). In many cases such approaches can be more successful in detecting the significant features. In particular, morphological segmentation on the height map often produces unsatisfactory results when there is a large flat surface with a sharp step, as shown in figure 2.15b. In such cases, there is no sharp ridge and regions are merged such that the boundary lies a significant distance from the transition. On the other hand, segmentation on the gradient map (figure 2.15c) is more satisfactory in such situations and places the feature boundary in the centre of the wall transition, which is a ridge of high gradient.

Other maps can also be used to segment the surface. Blunt and Xiao 2011 and Blunt and Scott 2012 have applied morphological segmentation to the surface after applying a LoG filter. This approach will detect the ridges of maximum curvature and has been successfully used to detect the dimpled structures of a laser textured hard disk drive.

In many ways, morphological segmentation is ideally suited to detecting texture units, as the segmentation approach can be interpreted as just looking for significant boundaries and determining regions based on the boundaries. Therefore, it is just a question of choosing a parameter such that the significant boundaries agree with the boundaries of the texture unit. This choice will depend on the surface in question. One method, proposed by Jiang and Whitehouse 2012 and Jiang et al. 2007, is to apply morphological segmentation to the ACF of the surface. Just as peaks are formed in the ACF due to the tile structure of the surface, the tile boundaries will lie on the valley lines of the ACF at minimal correlation. This approach has been shown to successfully identify the tiles for a range of surfaces.

2.6.3 Active contours

Another approach to segmentation is active contours (Kass et al. 1988; Caselles et al. 2001; Goldenberg et al. 2001; Chan and Vese 2001). Active contours start by placing a contour, C, on the surface and defining the contour energy, $E(C)$, based on properties of the surface. The contour is then evolved through space to try and minimise the energy. The final (minimum energy) contour then defines the boundary of the segmented object. Figure 2.16 shows the evolution of the algorithm on a test surface featuring a laser textured micro-dimple, with initial, intermediate and final contours shown. Clearly the definition of the contour energy will define where the boundary is placed. There are a wide range of energy functions that could be used. In image processing, two of the most common are known as geometric active contours (Caselles et al. 2001) and active contours without edges (Chan and Vese 2001). The geometric active contour is based on an edge detection approach such that

$$E = \alpha \int_0^1 |C'(s)|^2 ds + \beta \int_0^1 |C''(s)|^2 ds - \lambda \int_0^1 |\nabla z(C(s))| ds \tag{2.4}$$

where $C'(s)$ and $C''(s)$ are the first and second derivatives along the contour, α, β and λ are coefficients controlling the relative magnitude of each term, although commonly $\beta = 0$ is used, and $\nabla z(C(s))$ is the image gradient at that point. From equation 2.4 it is clear that, barring effects due to the contour shape, the energy will be minimised when the gradient around the contour is maximised.

An alternative approach, proposed by Chan and Vese 2001, is active contours without edges. This method is more similar to thresholding as described in section 2.6.1. Active contours without edges defines the contour energy as

$$\begin{aligned} E =& \mu.\text{Length}(C) + \nu.\text{Area}(inside(C)) + \lambda_1 \int_{\text{inside}(C)} |z(x,y) - c_1|^2 dxdy \\ &+ \lambda_2 \int_{\text{outside}(C)} |z(x,y) - c_2|^2 dxdy \end{aligned} \tag{2.5}$$

where the first two terms serve a similar purpose to the first two terms in equation 2.4 and similarly $\nu = 0$ is generally used. $z(x, y)$ is the surface height at the point (x, y) and c_1 and c_2 are the average heights inside and outside the contour respectively. λ_1 and λ_2 are constants controlling the position of the segmentation, with $\lambda_1 = \lambda_2 = 1$ being the standard choice. This approach minimises the variance between the two groups and is, therefore, similar to k-means clustering (Macqueen 1967) and Otsu's method (Otsu 1975), which also try to minimise the inter-region variance.

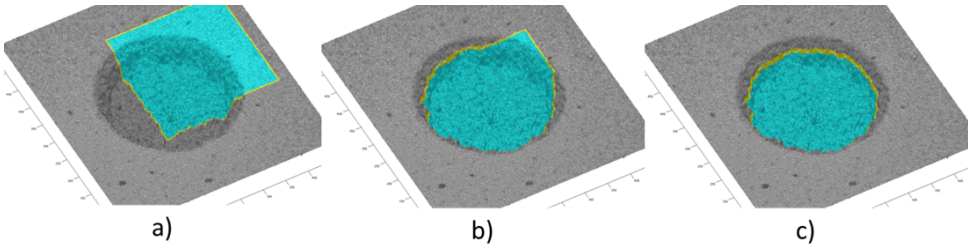

Figure 2.16: Evolution of the active contour algorithm showing: a) initial contour, b) intermediate contour and c) final contour. In this case, the active contours without edges algorithm was used.

One of the issues with active contours is the choice of initial contour. Ideally, this contour should be placed close to the optimal contour to ensure swift convergence. If it is not, then the solution may take a long time to converge, especially when the initial contour is far from the feature with little in the way of gradient for the contour to move down. In severe cases, the algorithm may become stuck in a local minima and not converge to the optimal solution.

Another issue is that if attempting to detect multiple features is that the numerical can increase significantly making the time for the algorithm to converge unreasonably long. Additionally, unless separate initial contours are created for each feature then the algorithm used must also allow for splitting the contour which is not always the case.

In surface metrology, there has been some limited use of active contours for surface segmentation. In particular Zhu 2012, Zhu et al. 2009a and Zhu et al. 2009b, have used active contours with the geometric active contour algorithm for segmentation of various structured surfaces, such as microprocessor chips and etched structures. In these cases, the authors appear to be able to accurately segment the features, although no quantitative tests are made.

2.7 Feature identification and determination of geometric attributes

Feature identification involves identifying the relevant geometric features on the surface, which are then used to calculate parameters. This section considers two methods of feature identification: via segmentation, using one of the methods in section 2.6, or via template matching, a technique to identify features without segmenting the surface. Additionally,

the process of determining relevant geometric attributes from these features is discussed.

2.7.1 Feature identification

The segmentation methods in section 2.6 partition the surface into different regions, but inherently assign feature and background or similar classifications. Feature identification aims to make these classifications. This can either be implemented as a subsequent step to segmentation, in which case the regions produced by segmentation are post-processed and analysed in order to identify the features and pattern units. Alternatively, template matching approaches can be used to identify the features and pattern units without segmenting the surface.

Feature identification based on segmentation

This family of techniques first involves post-processing the regions resulting from segmentation, for example for removing the ones that do not comply to specific size and shape criteria, or by merging regions into more relevant formations. Common approaches are to remove regions below a certain size (Senin et al. 2014) or that are an abnormal shape (e.g. not round) (Hartmann and Loderer 2013). Similarly, some segmentation methods (e.g. gradient thresholds) may result in regions that do not correspond to the feature being searched for, just to its boundary. In which case, if the region forming the boundary can be identified, any region enclosed by this boundary can be merged with it to obtain a filled region.

Another common issue with segmentation based feature identification is that regions intersecting the boundary of the image generally need to be excluded, as they may eventually result in features that are partially cropped out of the image, thus unusable for the determination of most geometric attributes. Finally, sometimes regions may need to be merged with each other before they can represent good candidates for feature identification. Whether such merging is required is dependent on the algorithm used, but is most common when using thresholding with multiple thresholds (Senin et al. 2007) or morphological segmentation (Senin et al. 2013). In some cases, having features consisting of multiple regions can be advantageous, for example if the combined regions provide a more accurate description of the feature than could be achieved by segmenting the feature into a single region. In such situations, these regions should be merged together to form a single feature. Suitable algorithms to perform such merging are often determined heuristically. A common approach is to mark the largest single region as the background

and merge all other adjacent regions together.

The classification of the post-processed regions can often be implied from the segmentation used, for example regions below a height threshold may be features. However, this is not always possible and some other heuristic may be needed to perform the final identification step. Typically, this identification is based on compliance to specific size, placement and form factor attributes, such as a dimple's circularity and size.

Feature identification based on template matching

Template matching is an alternative feature identification approach that does not rely on segmentation. Rather than segmenting the surface and using the resulting regions to identify the feature, template matching starts with a geometric model (template) of the surface feature and moves this model over the surface to search for a best match with the local topography, as an indication of successful feature identification. Deviations between the identified features and the template can then be assessed to determine the feature's properties.

Senin et al. 2011 describe template matching as consisting of three steps. Firstly, feature identification is performed. This step finds candidate regions in the image where features matching the template can be found. Secondly, feature extraction is performed, where the template is aligned with the feature in each candidate region and the region bounded by the template is extracted as a separate geometric entity. Finally, the nominal (aligned template) and measured (extracted region) geometry are compared to determine deviations from the model. Each of these steps will now be considered in more detail.

Feature identification is one of the key steps in template matching. Some method must be used to compare the template and surface to find points where there is a good agreement between the two. This provides the rough location of the features which can then be accurately matched to the template. Many methods have been developed to perform this feature identification. Primarily, these methods are based on finding a correlation between some property of the template and the surface.

Directly performing cross-correlation between the template and surface can be successful in some cases but is not ideal as the result is sensitive to the orientation of the template to the feature on the surface. Therefore, a variety of other properties of the template have been investigated for locating the features. Senin et al. 2011 use a method based on the ring projection transform (Lin and Chen 2008). Jiang et al. 2010 have also developed a similar technique called the structured region signature based on the point

signature method (Chua and Jarivs 1997). The ring projection approach looks at the sum of heights around rings of different radii, whereas the structured region signature approach takes a single radius ring. In either case, locations with a similar signal to the template indicate a possible match.

Once the features have been identified, coarse registration can be performed. This step aims to determine the approximate alignment of the feature so that the fine alignment step can be performed successfully. If the alignment is already known, for example if the template is rotationally invariant or is already determined by the feature identification step, then coarse alignment is not necessary. Alternatively, if the template consists of similar size area to the measured area, rather than just a single feature, then need for feature identification is lessened as coarse registration should cover all the significant features.

Various methods of coarse registration exist. The most simple approach to coarse registration, as used by Senin et al. 2011, is to try a fixed number of orientations and determine which has the best fit between the template and the model. Yu et al. 2011a have developed an alternative approach based on matching the salient points of the template and surface. This approach may be more complex, but is well suited to cases where feature identification has not been explicitly performed and the location of the features is not yet well known.

Fine registration is used to create an accurate alignment between the template and surface. The most widely used algorithm for fine registration is the iterative closest point (ICP) method (Besl and McKay 1992). This algorithm computes the rigid transformation of the template which minimises the least squares sum of differences between the template and surface. However, being an iterative algorithm it is relatively computationally expensive and can easily become stuck in local minima. To avoid this issue, good coarse registration is needed.

Once the surface and template are well aligned, the difference of the two can be taken to determine form error in the surface model. This error could be assessed in a number of ways depending on the surface, either using the sum of errors over the surface (Lin and Chen 2008) or conventional dimensional tolerancing approaches, such as tolerance zones (ISO 2013).

While template matching is a powerful tool that identifies features on a surface without relying on segmentation, it has some drawbacks. Firstly, template matching relies on having a template that is representative on the nominal or expected surface, ideally this would be obtained from a CAD model or similar. However, it is not always possi-

ble to obtain such a template, especially in cases where the surface is poorly specified. Additionally, template matching does not directly extract the geometrical properties of the surface, only their difference from the template. In many cases, this is not a problem. However, large systematic errors due to misspecification or similar can swamp other errors, which may be of more interest and make analysis difficult.

The above sections have described two methods of feature identification. Feature identification based on segmentation follows directly from the segmentation methods discussed in section 2.6 and is primarily aimed at removing noise and misclassification from the segmentation and identifying regions as feature or background. Template matching is an alternative approach, which relies on knowledge of the features on the surface to determine the regions on the surface which best match. While template matching does not explicitly segment the surface, a segmentation can often be implied from the identified features.

2.7.2 Geometric attribute determination

The final step in determining feature parameters is to determine the relevant attributes of the individual features. The statistics of these attributes, computed from a sample of features collected while inspecting the measured area, are what constitutes the feature parameter. There are a vast array of different attributes that could be considered. It is, therefore, useful to consider attributes of the individual features and attributes of the pattern separately.

Attributes of the feature refer to properties of the individual features such as their shape and size. In general, attributes of the features will either fit a shape to the boundary to determine its size (e.g. fitting a circle to determine feature radius) or will use some or all of the height data in the feature to determine the attribute (e.g. roughness or depth of the feature).

Attributes of the pattern, on the other hand, focus on the properties of the lattice which makes up the structured surface (e.g. the spacing and angle between features). To determine these attributes, a consistent reference point is needed for each feature or tile to allow for consistent calculation of the parameters. For circular shaped features, such as the dimpled surface test case, one common approach, as applied by Kong et al. 2010, is to fit a circle (or other appropriate shape) to the boundary of the feature and use the centre of the fitted shape as the centre of the feature. Other similar approaches could be used based on placing the centre at the centre of mass of the feature.

Once the relevant attributes have been computed, feature parameters can be calculated by calculating the relevant statistics of the attribute distribution. Commonly used statistics are the mean, standard deviation and median. However, a wide range of other statistics could be used, such as, mean of the five smallest or largest values.

2.8 Assessment of measurement uncertainty

In order for a measurement to be complete, it must have an associated measurement uncertainty, which describes the range in which the true value of a measured property is believed to lie. Therefore it is important to be able to estimate the uncertainty in the measured properties of a structured surface, as the uncertainty can have a significant impact on tolerances and the significance of any correlations drawn from the measurement results. For structured surfaces, uncertainty has not previously been widely considered. Therefore, methods to assess the uncertainty in structured surfaces should be developed to allow for full characterisation of such surfaces.

The main document guiding the assessment is the Guide to the expression of uncertainty in measurement (GUM) (JCGM 2008a). An updated version of the GUM is currently in committee draft (JCGM 2015). This section reviews the principles of measurement uncertainty as they are set out in the current version of the GUM (JCGM 2008a). Cases where there are major differences in the new draft are also noted.

2.8.1 Uncertainty terminology and definitions

There are lots of terminology and definitions associated with determining uncertainty, many of which are often used incorrectly or inconsistently. The official definitions of many of these terms are set of in the International vocabulary of metrology (VIM) (JCGM 2012) as well as their usage in the GUM (JCGM 2008a). This section provides a brief overview of some of the important terminology and their definitions.

- Accuracy - describes the difference between a measured value and the true value of that quantity. While related accuracy does not directly describe measurement uncertainty as it only refers to a single value, whereas the uncertainty the expected value. Additionally, it is rarely possible to determine the accuracy exactly as there will almost always be some uncertainty in what the true value is.

- Trueness - describes the difference between the expectation of a measurement and

the true value. This is the inverse idea to systematic measurement error. Just as for accuracy it is difficult to determine what the true value actually is.

- Measurement error - is the difference between a measurement and a reference value. This reference value may be the supposed true value or the result of a previous measurement or calibration.

- Systematic measurement error/Bias - is the component of measurement error that remains constant between measurements. Therefore, it does not introduce any uncertainty into the measurements, but may result in bias in measurements which may be corrected if appropriate. It is normally defined as equal to the expectation of the measurement error.

- Random measurement error - is the component of measurement error that varies between measurements and is what makes up the uncertainty. By convention, it is normally defined with expectation 0.

- Precision - is the closeness of repeated measurements under the same or similar conditions. This is normally what is mean by the uncertainty of a measurement and is usually defined by the standard deviation of a set of measurements, although other metrics of precision could also be used.

- Measurement repeatability - Is the measurement precision while the measurement conditions (or strictly a defined set of measurement conditions) are kept constant or as close to constant as possible. The relevant measurement conditions could include measurement position or instrument settings such as brightness.

- Measurement reproducibility - Is the measurement precision while a subset of the measurement conditions are allowed to vary. This may involve repositioning the sample or adjusting measurement positions and will normally result in a higher uncertainty than that given by the measurement repeatability and can be considered more representative of real measurement results.

2.8.2 Sources and types of uncertainty

Measurement uncertainty describes the expected dispersion of a set of measurement results. There can be many sources of error in an measurement set-up such as, electronic noise in the detector, temperature fluctuations or repeatability of the positioning system.

Many of these error sources will be random, giving varying values between measurements or over time. Such random errors will result in uncertainty, as multiple measurements of the same thing will result in different values.

Additionally there may be some systematic errors, which cause a constant error between the measured and 'true' value. Such errors will result in a bias in the measurement. Where possible bias should be corrected, such that the mean of multiple measurements agrees with the 'true' value of the quantity being measured. In reality, the 'true' value can never be known entirely accurately and so there will be some residual bias which remains unknown. This can be estimated and considered another contribution to the uncertainty.

The GUM (JCGM 2008a) breaks the sources of uncertainty into two categories, which are not related to the source but to how the uncertainty is determined. Type A uncertainties are those determined by statistical analysis of a series of observations, whereas type B are determined using a priori knowledge or other means. In practise, these differences only affect how the initial uncertainty estimate is reached.

2.8.3 Evaluation of uncertainty

A measurement and its uncertainty can be fully described by the probability distribution function (PDF) of the property being measured. However, in reality this is not a practical way to describe a measurement and uncertainty is normally described by a confidence interval, giving a range within which the measurement will fall with a certain probability. Combined with a description of the shape of the PDF, the confidence interval is sufficient to describe a measurement.

For type A uncertainties, it is assumed that a set of n repeat measurements are representative of the underlying distribution. The standard deviation of the measurements given by

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2} \tag{2.6}$$

where, x_i is the ith measurement and μ is the mean of the measured values, gives a basis to define a confidence interval.

The standard deviation in equation 2.6 can be used directly to estimate uncertainty. However, such an uncertainty implies that only one actual measurement of the quantity of interest is made. While in some cases this may be true, the standard approach for a type A uncertainty is to use the mean of the repeat measurements as the best estimate of the value being measured. In this case the mean of the results is the parameter of interest

and the standard deviation of the mean determines the uncertainty given by

$$u = \frac{\sigma}{\sqrt{n}} \tag{2.7}$$

where the standard deviation of the mean here is defined as the standard uncertainty, u. Note, the current committee draft for the new version of GUM (JCGM 2015) recommends using

$$u = \sqrt{\frac{n-1}{n-3}}\frac{\sigma}{\sqrt{n}} \tag{2.8}$$

to determine standard uncertainty. This modification accounts for the lack of knowledge of the true standard deviation when few measurements are made. However, this difference is not significant for this thesis.

For type B uncertainties, u is determined via prior knowledge or other, non-statistical means, such as the value from a measurement certificate.

2.8.4 Uncertainty distributions

As well as the value of u, it is also important to know the shape of the PDF. While, in theory, an arbitrary PDF could be used to describe the uncertainty distribution, for most practical cases the PDF is approximated by a well defined function to simplify the analysis. Several standard functions are used which have useful properties and provide a good approximation for most processes. Here, some of these useful distributions are considered in more detail: the Gaussian, uniform, triangular, curvilinear trapezoid and Student's t-distributions.

Gaussian distribution

The Gaussian (also known as normal) distribution is probably the most commonly used distribution and is a good model for variations occurring in many processes. Due to the Central Limit Theory, a combination of multiple different distributions will tend towards a Gaussian distribution (Hughes and Hase 2010). This means that many complex processes will result in roughly Gaussian distributions. The PDF of a Gaussian is defined as

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.9}$$

the shape of this distribution as a function of σ is shown in figure 2.17(a).

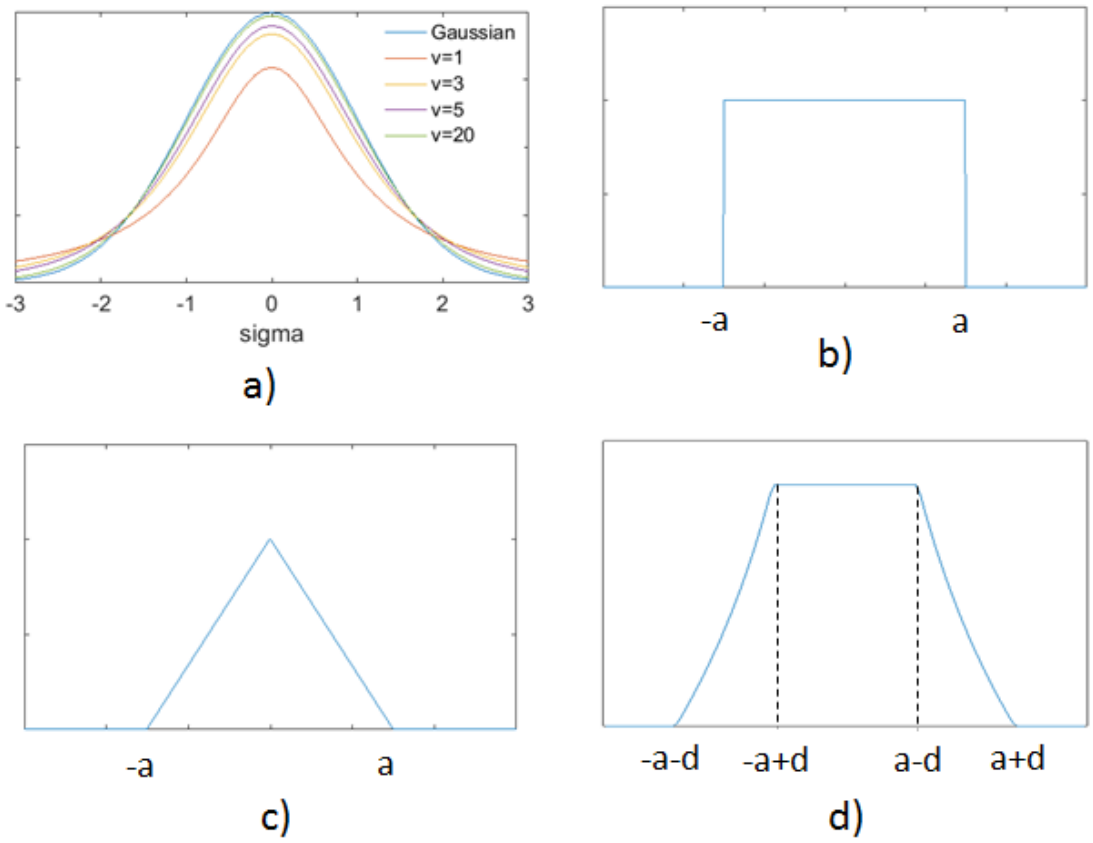

Figure 2.17: Plots of PDFs of different uncertainty distributions. a) Gaussian and t-distribution with a range of degrees of freedom, v. b) Rectangular distribution. c) Triangular distribution. d) Curvilinear Trapezoid

Student's t-distribution

The Student's t-distribution can be viewed as a modified Gaussian distribution which is more applicable for small sample sizes. The shape of the distribution, as shown in figure 2.17(a), is controlled by the degrees of freedom, v. For independent, repeated measurements the degrees of freedom are determined by the number of measurements, such that for n measurements

$$v = n - 1 \tag{2.10}$$

As can be seen from figure 2.17(a), the difference between the t-distribution and the normal distribution is greater for small v, whereas for large v the solutions converge. Indeed, the normal distribution is just the limit as $v \to \infty$.

Uniform distribution

Another often used distribution is the uniform (or rectangular) distribution. A uniform distribution has an equal chance of returning any value within a certain range and is described by the PDF

$$\begin{aligned} P(x) &= \frac{1}{2a} && \mu - a < x < \mu + a \\ P(x) &= 0 && \text{otherwise} \end{aligned} \tag{2.11}$$

where $2a$ is the width of the distribution. The standard deviation of this distribution is then

$$\sigma_U = \frac{a}{\sqrt{3}} = \frac{2a}{\sqrt{12}} \tag{2.12}$$

An example of a rectangular distribution is shown in figure 2.17(b). Generally, a uniform distribution is a good model for situations where maximum and minimum values of distribution are known, but there is no extra knowledge about what values are most likely (or all values are considered equally likely).

Triangular distribution

The triangular distribution is similar to the uniform distribution, but triangular in shape rather than rectangular. Therefore, it has a PDF of

$$\begin{aligned} P(x) &= \left| \frac{x - a}{a^2} \right| && \mu - a < x < \mu + a \\ P(x) &= 0 && \text{otherwise} \end{aligned} \tag{2.13}$$

where again the width is $2a$. This gives the standard deviation as

$$\sigma_{Tri} = \frac{a}{\sqrt{6}} \tag{2.14}$$

An example of a triangular distribution is shown in figure 2.17(c). Such a distribution is often used in situations where the bounds are known and it is expected that values close to the mean are more likely than values far from the mean.

Curvilinear trapezoid distribution

In some cases a uniform distribution may be desirable, but the appropriate limits are uncertain. In such situations a curvilinear trapezoid can be used to account for this uncertainty JCGM 2008b. The curvilinear trapezoid centred on zero, $\mathrm{CTrap}(-a, a, d)$ has

nominal bounds on the rectangular distribution of $\pm a$ and uncertainty in these bounds of $\pm d$, resulting in a PDF of

$$\begin{aligned} P(x) &= \frac{1}{4d}\ln\left(\frac{a+d}{-x}\right) && -a-d < x < -a+d \\ P(x) &= \frac{1}{4d}\ln\left(\frac{a+d}{a-d}\right) && -a+d < x < a-d \\ P(x) &= \frac{1}{4d}\ln\left(\frac{a+d}{x}\right) && a-d < x < a+d \\ P(x) &= 0 && \text{Otherwise} \end{aligned} \tag{2.15}$$

The standard deviation of this distribution is then

$$\sigma_{CTrap} = \sqrt{\frac{4a^2}{12} + \frac{d^2}{9}} \tag{2.16}$$

An example of a curvilinear distribution is shown in figure 2.17(d).

2.8.5 Propagation of uncertainty

As mentioned in section 2.8.2, there are many sources of uncertainty that can contribute to the final measurement. Often it is more practical to determine the uncertainty of each of these sources separately than to try and measure the entire uncertainty of a measurement directly. Therefore, it is necessary to be able to understand how these individual uncertainties affect the final result and how to propagate them into the combined uncertainty of the entire measurement.

In general, there exists an equation $y = f(x_1, x_2, ...x_n)$, to combine the n individual inputs into the measurement result, y. For example, to determine the combined length of two rods the equation would be

$$L_T = L_1 + L_2 \tag{2.17}$$

where L_1 and L_2 are the lengths of the rods and L_T is the combined length. The standard approach is then to use a first order approximation to the Taylor expansion to calculate the combined standard uncertainty of the measurements given by the law of propagation of uncertainty

$$u_c^2 = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i}\right)^2 u_i^2 \tag{2.18}$$

where u_c is the combined uncertainty, x_i is the ith input and u_i is its associated uncertainty. Equation 2.18 assumes that the inputs are all uncorrelated. If this is not the

case, the equation must be modified as discussed in section 2.8.7. A full derivation of equation 2.18 can be found in the GUM (JCGM 2008a).

The law of propagation of uncertainty, given in equation 2.18, relies on several significant assumptions. Firstly, that higher order terms in the Taylor expansion terms are negligible. For all cases in this thesis, higher order terms will not be a major concern. However, if they are, extra terms can be added to the law of propagation of uncertainty to account for them (JCGM 2008a). Additionally, it is assumed that the Central Limit Theorem can be applied and the combined distribution is approximately Gaussian. While this assumption tends to be valid if the number of inputs is large, care should be taken when the number of inputs is relatively low. In general, the validity of the Central Limit Theorem will not only depend on the number of inputs, but also their distributions, so no universal limit can be made. Thirdly, it is required that f is well known and everywhere differentiable. For many cases in surface metrology, f may not be differentiable or may not be easy to evaluate. In such cases, further approximations may be made. The derivative $\frac{\partial f}{\partial x_i}$ may be replaced with a more general sensitivity coefficient c_i which can be determined experimentally or by theory. Alternatively, a numerical approximation of the gradient may be made such that

$$\frac{\partial f}{\partial x_i} u_i \simeq \frac{1}{2}[f(x_1, ...x_i + u_i, ...x_n) - f(x_1, ...x_i - u_i, ...x_n)] \tag{2.19}$$

This approximation assumes that the underlying function is approximately linear in the region considered, which is not always the case.

When the second and third assumptions above do not hold, one option is to determine the uncertainty numerically using Monte Carlo techniques. This approach can be an effective way to determine uncertainty even when f cannot be defined algebraically or the resulting distribution is not Gaussian, and is discussed in more detail in section 2.8.8.

2.8.6 Expanded uncertainty

While knowledge of the combined uncertainty and the uncertainty distribution of a system is sufficient to fully describe the uncertainty, in many cases it is desirable to give a uncertainty value which better describes the probability of finding the true value in a given range. This description is given by the expanded uncertainty, which is determined by multiplying the combined uncertainty by a coverage factor, k. This expanded uncertainty is what is generally quoted when the uncertainty is stated, in the form

$$Y = y \pm k u_c \tag{2.20}$$

The value of the coverage factor defines the number of standard deviations the confidence interval covers. However, how this value relates to a percentage coverage of the PDF will also depend on the shape of the PDF. Often k is chosen such that it corresponds with the 95 % confidence level, but other percentages can be used if appropriate. 95 % confidence corresponds to a value of $k = 1.96$ for a Gaussian distribution (this is often approximated by $k = 2$). However, for a uniform distribution the same confidence level corresponds to a value of $k = 1.68$. This difference highlights the importance of considering the shape of the distribution when calculating the expanded uncertainty.

2.8.7 Correlation

The law of propagation of uncertainty in equation 2.18 considers that the variables are independent of each other. In many situations this is a reasonable assumption. However, for some cases there may be significant correlation between variables which should be accounted for. In the general case, correlation between variables can be considered by determining the correlation coefficient between all pairs of variables. The correlation coefficient between two variables, $r(x_i, x_j)$ is given by

$$r(x_i, x_j) = \frac{E\left((x_i - \mu(x_i))(x_j - \mu(x_j))\right)}{\sigma_i \sigma_j} \tag{2.21}$$

which should lie in the range ± 1. Once the correlation coefficient is known, then the combined uncertainty of the measurement is given by (see JCGM 2008a)

$$u_c^2 = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}^2 u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u_i u_j r(x_i, x_j) \tag{2.22}$$

This equation can also be written in matrix form, which for many situations may be more convenient, as

$$u_c^2 = c \Sigma c^T \tag{2.23}$$

where Σ is the co-variance matrix, with $\Sigma_{ij} = E\left((x_i - \mu(x_i))(x_j - \mu(x_j))\right)$, and c is a row vector of the sensitivity coefficients, $\frac{\partial f}{\partial x_i}$.

2.8.8 Monte Carlo methods

Equation 2.18, relies on several assumptions. Mainly that the combined uncertainty has a Gaussian distribution due to the Central Limit Theorem and that the function propagating the uncertainty is everywhere differentiable. These assumptions are not always met.

In some situations analytical solutions may be possible, for example when combining two rectangular distributions (Dietrich 1991), but analytical solutions become increasingly complicated as the problem difficulty increases. For cases where the analytical solution does not exist or is too difficult to solve, Monte Carlo methods (JCGM 2008b; Schwenke et al. 2000; Lemieux 2009) provide an alternative approach based on statistical sampling.

Monte Carlo methods are a class of technique that rely on random sampling to estimate the output. The basic principle is that given a function $y = f(x_1, x_2, ...x_n)$, the distribution of y can be estimated by repeatedly sampling the function using random values of x_i, sampled according to the PDF of each x_i. Each set of random numbers will provide a single estimate of y. After many repeats, these estimates of y will build up to provide an estimate of the PDF of y. This PDF can either be used directly to estimate uncertainty or used to estimate a model of the distribution. Figure 2.18 shows an overview of this approach.

The advantage of the Monte Carlo method is that it is not dependant on the shape of the underlying or final distributions and does not require differentiation of f. However, due to the statistical nature of Monte Carlo methods, a large number of samples are needed to get an accurate result. Generally, there is no simple relationship between number of samples and accuracy of the result. Indeed, due to the random nature of Monte Carlo methods no finite number of samples can be guaranteed to give a desired accuracy (JCGM 2008b). Generally, several thousand samples at least are required, with the main limiting factor generally being computation time. An alternative approach, to ensure results are within a given tolerance would be to use an adaptive method (JCGM 2008b), which increases the number of samples used until the results converge sufficiently. Therefore, the main disadvantage of the Monte Carlo approach is that due to the large number of samples required it is relatively computationally intensive, at least compared to the approach in equation 2.18.

2.8.9 Application of measurement uncertainty to structured surfaces

Previous work in the field has only considered ways to characterise structured surfaces and their geometric properties, no detailed consideration has previously been given to how to determine the associated uncertainties in such measurements. Therefore, assessment of uncertainty is an important area of work, worthy of further research. Due to the lack of previous work in the same area it may be of interest to consider related areas and how

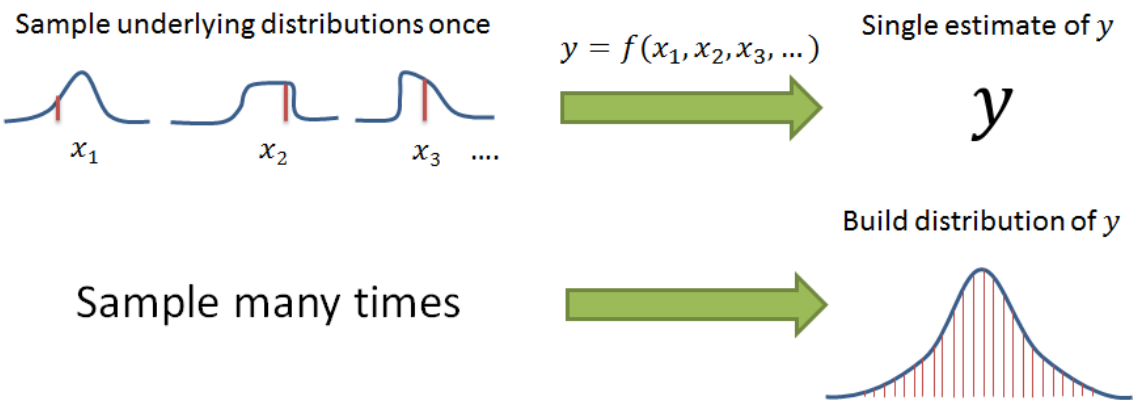

Figure 2.18: Overview of Monte Carlo algorithm. A single estimate of y can be made by sampling from the underlying distributions. By sampling many times, many estimates of y are made and the distribution of y can be estimated.

they assess uncertainty.

The obvious area to look for related work on uncertainty estimation is in computer vision and image processing where many of the segmentation and analysis techniques used for structured surfaces originated. A limited amount of research has considered the uncertainty due to segmentation processes. De Santo et al. 2004 have propagated image measurement uncertainty through the calculation of the Gaussian and gradient of the image and applied these uncertainties to the calculation of lengths and areas of segmented images. Similarly, De Santo et al. 2000 and Anchini et al. 2007 have calculated analytical uncertainties in various edge detection operators. These approaches consider the uncertainty of each pixel in the field of view and propagate them through to the final measurement uncertainty. Several simplified models must be used to avoid issues caused by non differentiable functions and relatively large correlation matrices are required to account for relationships between nearby pixels. Despite these limitations these approaches seem to be successful in determining an uncertainty in the geometric properties of interest. However, only relatively simple operators are considered and it is not clear how easily such an approach would generalise to more complex segmentation methods.

Another area to investigate is how uncertainties are determined for surface metrology of stochastic surfaces. Assessment of uncertainty in field parameters is difficult and rarely performed as many influence factors of the instrument can effect uncertainty (Leach 2011; Harris et al. 2010). However, uncertainty can be assessed using the instrument's metrological characteristics (Haitjema 1998; Haitjema 2013). The metrological characteristics of a

surface texture measuring instrument convert the many influence factors that can effect an instrument's uncertainty, such as scan speed, illumination intensity, stage repeatability, etc. into a set of characteristics that describe the geometrical errors caused by these influence factors and can be directly measured. The metrological characteristics and how to measure them are described in Giusca et al. 2012a; Giusca et al. 2012b; Giusca and Leach 2013.

Metrological characteristics appear to be a natural way to describe the uncertainty created by the instrument as they describe observed errors in an easily measurable way. However, care is required when propagating the metrological characteristics into uncertainty of structured surfaces as there is no simple relationship between the two. This suggests that Monte Carlo simulation will be needed, at least initially to avoid making false assumptions.

Finally, the simplest way to estimate the uncertainty may be via repeated measurement of a feature, i.e. direct type A uncertainty evaluation. However, the difficulty in such a situation is to consider the full range of conditions observed in a real measurement, while not requiring an excessive number of measurements.

2.9 Summary

This chapter has provided an overview of the current state of the art and main factors to consider when measuring and characterising structured surfaces. The main points are:

1. A wide range of instruments capable of measuring structured surfaces exist and the most suitable instrument will depend on the surface being measured and the desired resolution.

2. Field parameters are well developed from characterisation of conventional surface texture. However, field parameters also have many limitations for characterising structured surfaces and generally do not account for the deterministic tile layout and shape, typical of structured surfaces. Therefore, they may be insensitive to changes in topography, which change the functional performance of the surface. For this reason, field parameters are often poorly suited to characterising structured surfaces. However, in situations where functional relationships with field parameters can be established or only verification is required, field parameters may be suitable due to their relative simplicity and widespread understanding.

3. Segmentation-based methods show greater promise for characterisation of structured surfaces as they allow for characterisation of the geometric properties of features, which are closely related to how such surfaces are designed. Such parameters allow for easier direct comparison with the specification and are more likely to correlate with functional performance. For these reasons, this thesis will primarily focus on the use of segmentation techniques. However, such techniques are complicated with a large number of control parameters to optimise. Therefore, chapter 4 will investigate how various segmentation methods respond to changes in their control parameters and the differences between segmentation algorithms.

4. The principles behind several types of segmentation algorithm: thresholding, morphological segmentation and active contours, that have previously been used for the analysis of structured surfaces, were considered in detail and pros and cons of each method was discussed. These methods are the main methods used throughout the rest of this thesis.

5. Template matching approaches are an alternative approach to characterising surfaces, which rely on having a relevant template of the features from a CAD model or similar and determine errors in this model rather than directly characterising the surface. For the surfaces used in this thesis a suitable template was not easily available and the main interest was in characterising the geometric properties. Therefore, template matching approaches are not considered further.

6. Estimation of uncertainty in structured surfaces is an area where little research has been done and further work is of interest. Therefore, the fundamental principles of uncertainty estimation where reviewed. These ideas will form the mathematical basis to allow for the assessment of uncertainty in the geometrical properties of structured surfaces in this thesis.

7. Several possible approaches to assess uncertainty in structured surfaces were identified. The most simple way being a reproducibility analysis using repeated measurements. However, suitable experiments must be designed and performed to account for the wide range of factors that can vary between measurements. An alternative approach would be to calculate uncertainties analytically based on knowledge of the uncertainties introduced by the measurement instrument. Such approaches have previously been used in image processing (De Santo et al. 2004; De Santo et al. 2000; Anchini et al. 2007), but are currently limited to easily analysable functions

with well understood measurement noise, which would limit their usefulness for surface measurement. A third approach would be to use the instrument's metrological characteristics to determine the uncertainty in a similar way as for areal field parameters (Haitjema 2013; Haitjema 1998). However, new techniques would need to be developed to propagate the metrological characteristics into the relevant geometric properties and their uncertainties.

8. Based on this assessment of the literature, the remainder of this thesis will focus on segmentation based techniques to characterise structured surfaces. These approaches show the greatest promise to produce characterisations that are functionally relevant. In particular, the thesis will investigate the difference between the different segmentation methods and how changing various control parameters affects the results as it is important to understand how these factors affect the characterisation. Additionally, uncertainty for structured surfaces is still a relatively unexplored topic and the thesis will also investigate methods to assess uncertainty in structured surfaces.

Chapter 3

Measurement and characterisation set-up

This chapter presents the various samples, instruments and analytic techniques that are used in the rest of the thesis. Where appropriate, consideration will also be given to advantages and disadvantages of particular measurement equipment and other important factors.

3.1 Samples used for measurements

The samples used in this thesis are silicon nitride disks, with a repeating pattern of features manufactured by laser texturing. The features, known as dimples, are nominally cylindrical pits with a specified diameter and depth. Figure 3.1 shows one of these disks with dimples at a range of scales. For this thesis disks with dimples of nominal diameter 50, 150 and 300 μm were used. The dimples had 10 μm nominal depth and 20 % nominal surface area density.

The disks where designed to demonstrate friction reduction in sliding contacts. Under hydrodynamic lubrication the shape of the dimples causes additional lift to be generated between the two surfaces reducing friction (Etsion et al. 1999; Wang 2014). Tests on these disks have been carried out by the materials group at NPL and Southampton NCATS. However, their results were inconclusive, and resulted in higher friction under the experimental conditions used. However, further research into the functional performance of the disks is ongoing and it is believed that friction can be reduced under suitable experimental conditions.

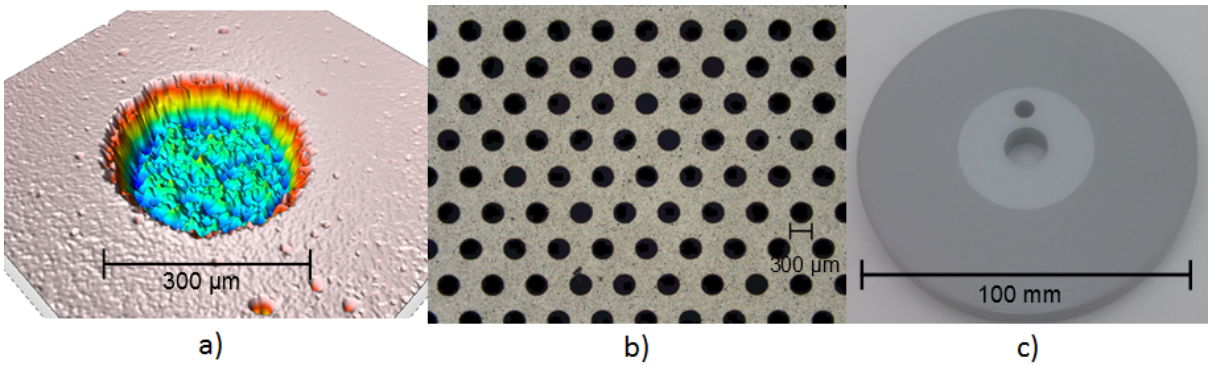

Figure 3.1: Examples of dimple features at different scales: a) 3D heightmap of a single dimple. b) Microscope image of an area of the disk, approximate field of view 5 mm×4 mm. *c) Image of the entire disk.*

3.2 Measurement instruments used

This section considers the instruments used for measurements in this thesis. For the measurement of structured surfaces, of the size used in this thesis, 3D optical microscopes are ideally suited due to their relatively fast measurement speed, compared to stylus based methods, and suitable measurement range, with the field of view varying from one to several features , depending on the precise set-up used. Two different instruments are used here: laser scanning confocal microscopy and focus variation microscopy. Both these approaches are well suited to measuring the relatively high slopes often found in structured surfaces. This section discusses the operating principles of these instruments in more detail. An overview of the general principles of a range of instruments has been presented in section 2.4.1.

3.2.1 Laser scanning confocal microscope

One of the main types of instruments used for in this thesis is laser scanning confocal microscopy, which relies on the confocal (or pinhole) effect to determine the height of the surface, as shown in figure 3.2. A monochromatic laser is focused on a point in space. The laser light reflected from the surface travels to the detector which has a pinhole just in front of it. The system is aligned such that if the light is reflected from the focal point it will be in focus at the pinhole and detector and a high intensity will be recorded. If the surface does not lie on the focal point then the reflected light will not be in focus on the pinhole. In this case, most of the light will be blocked by the pinhole and only a very

low intensity will be recorded. By scanning the system in the z direction the height of the surface can be determined by finding the maximum intensity. It is important to note that the laser focuses on a point on the surface and must scan in the xy plane to measure the entire surface. Various methods are possible to achieve such a scan, either using a spinning Nipkow disk which has holes in a specific pattern to focus light on different parts of the surface, using a series of deformable or movable mirrors to adjust the focus or by physically moving the sample, although the last approach is rarely used in commercial instruments due to its slowness (Artigas 2011).

Laser scanning confocal microscopes are capable of reasonably high measurement speed, and good lateral and vertical resolutions, on the order of 1 µm and 10 nm respectively. Additionally, they generally use higher numerical aperture (NA) lenses than comparable interferometric instruments, which allows for measurement of higher slopes than would be possible using a comparable white light interferometer. This is an important property for the measurement of structured surfaces, which often contain significant slopes that are important to measure.

The particular instrument used is an Olympus Lext OLS4000 (*Olympus LEXT OLS4000*). Primarily the 20× lens was used. The 20× has NA 0.60, vertical resolution 10 nm and a 0.64 × 0.64 mm field of view. 50× and 100× lenses were also available but the 300 µm diameter disks do not fit in the field of view for these lenses.

3.2.2 Focus variation microscope

The other main type of instrument used is the focus variation microscope. The set-up is essentially a simple optical microscope consisting of a white light source and detector, as shown in figure 3.3. The system scans in the z direction and calculates the sharpness of each pixel at each step. Sharpness can be calculated using a region around each pixel. Many definitions of image sharpness are possible and it is unknown what algorithm is actually used in the instrument. One possible definition, which will serve to illustrate the principles, is that sharpness can be defined as the standard deviation in the window

$$F(x,y,z) = \sqrt{\frac{1}{N_{win}} \sum_{win} (I(x,y,z) - \mu_{win,z})^2} \tag{3.1}$$

where $F(x,y,z)$ is the sharpness at that height and position, N_{win} is size of the window, $I(x,y,z)$ is the intensity at that height and position and $\mu_{win,z}$ is the average intensity over the window at that height. Considering the sharpness of a single pixel will result in

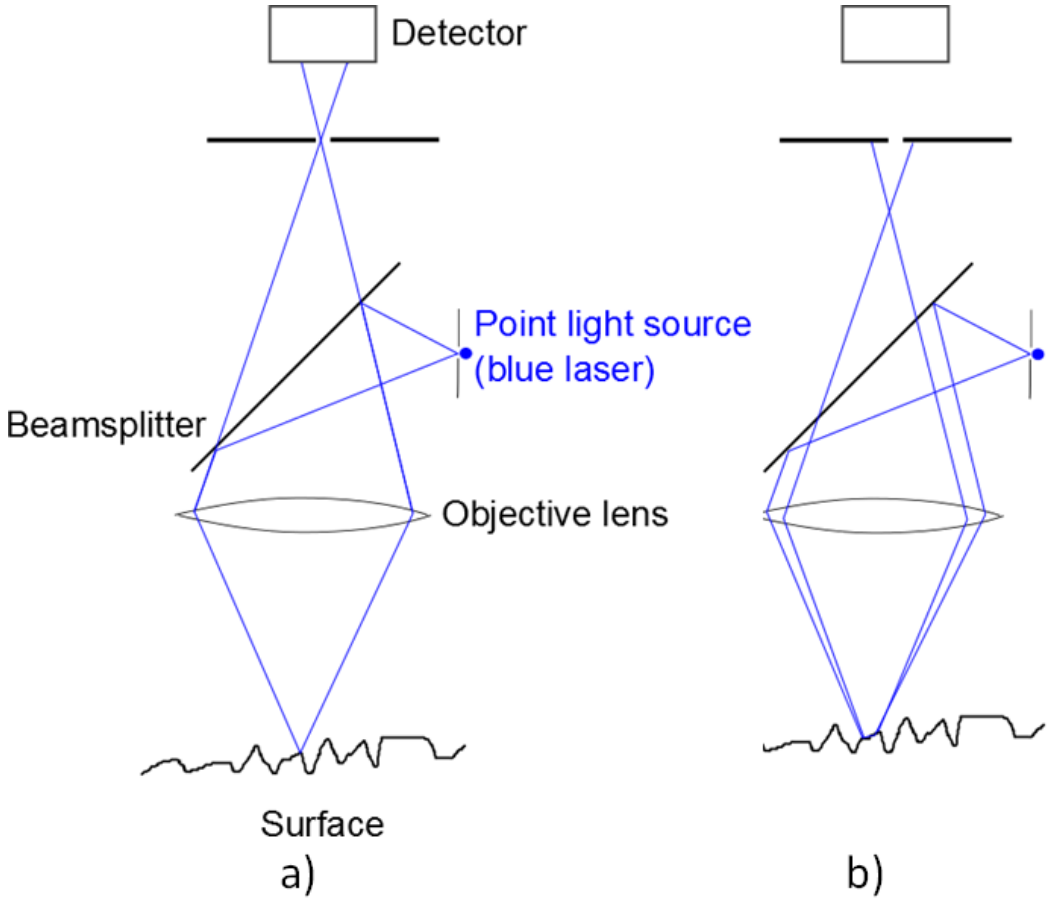

Figure 3.2: The principles of a scanning confocal microscope. a) When the distance between surface and objective lens is correct the laser is focused on the surface. Reflected light then passes through the pinhole and a strong signal is detected. b) If the height of the surface changes The light is not focussed on the surface and most of the reflection will be blocked by the pinhole. In this case only a weak or no signal is detected. By scanning in z the surface height is determined by finding the peak in the signal. The microscope scans the xy plane to measure different points on the surface.

a curve of sharpness against height. The height of each pixel is determined by finding the height at which the pixel is in focus and has maximum local sharpness (Helmli 2011).

As the focus variation microscope uses a white light source and conventional imaging system it does not rely on specular reflection to determine the surface in the same way as the laser confocal microscope or white light interferometers do. This allows the focus variation microscope to measure steep slopes with relatively low NA lenses by using diffuse reflections from the surface. The low NA of the lenses also allows for a larger

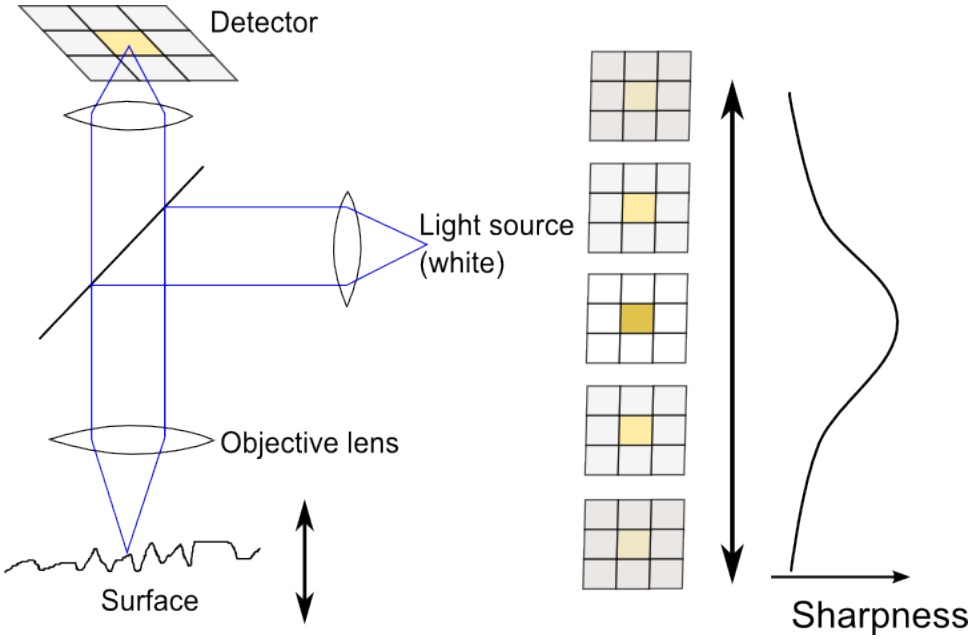

Figure 3.3: The principles of a focus variation microscope. Light is reflected from the surface before being collected by the detector, much like a standard microscope. As the surface is scanned an area around each pixel is used to determine the sharpness. The height of a pixel is determined by finding the maximum sharpness.

stand-off distance than other instruments capable of measuring such slopes. The disadvantage of this set-up is that it is not possible to measure very smooth surfaces where the sharpness has no clear peak. Additionally, to determine the sharpness, a moving window operation must be performed on each image. This window limits the lateral resolution of the instrument, which is generally several times what is achievable by other equivalent techniques.

The focus variation instrument used in this work is an Alicona Infinite Focus G4 (*Alicona Infinite Focus*). As for the laser confocal, primarily the 20× lens was used, which has NA 0.40, vertical resolution 50 nm and a 0.715 × 0.544 mm field of view. Higher magnification lenses were available, but again the 300 µm diameter disks do not fit in the field of view for these lenses.

3.3 Characterisation techniques and algorithms used for analysis

Four different segmentation techniques: height thresholding, gradient thresholding, morphological segmentation and active contours are used for a variety of purposes throughout this thesis. The general principles of these segmentation techniques have already been described in chapter 2 and are not repeated here. This section provides an overview of the algorithms used along with some important specific points. Unless otherwise noted, the algorithms were programmed using Matlab.

3.3.1 Pre-processing

Pre-processing consists of two main steps: levelling and filtering. Levelling was performed by subtraction of a least squares plane. The approach used allows exclusion of points, determined by a binary mask when fitting the plane. This approach makes it possible to exclude the feature, which would otherwise skew the results, from the fitting. However, knowledge of the feature position is required to exclude it. Therefore, the segmentation algorithm is generally run multiple times. First on unlevelled data, or data levelled with all points, to provide an initial segmentation. This segmentation provides the mask for subsequent levelling. In general, it was found that two or three levelling steps were sufficient to produce a well levelled surface. However, care must be taken to ensure that the initial segmentation produces a sensible result, otherwise the levelling will fail. The necessary steps depend on the measured data and segmentation method. If the measured data is reasonably level, then it may be possible to perform levelling without any additional preprocessing. However, for the height threshold in particular, it is also necessary to ensure that the mean height is near zero, as any change will effectively modify the threshold used. Therefore, the mean height is subtracted from the surface to ensure that the mean is zero.

Filtering was performed using a Gaussian filter with edge effects corrected (Seewig 2005; Seewig 2013). Although, in most cases the filter size is small enough that the edge correction is insignificant, as long as the feature is reasonably central. The filter size can either be specified in terms of a standard deviation of the Gaussian kernel or a cut-off where the frequency response drops to 50 %. It is straightforward to convert between these two descriptions, as described in section 2.4.2. Unless otherwise noted, a filter of $\sigma = 4$ pixels was used. This corresponds to a cut-off of 9.3 µm using the focus variation

microscope and 13.3 µm using the confocal microscope.

3.3.2 Height thresholding

The height threshold is the most straightforward of the segmentation algorithms used. It produces a binary mask where points with heights below the threshold value are 1 and points above are 0. As noted previously this, and all other segmentation algorithms, should be run at least twice to ensure that the surface is properly levelled. For the height threshold, it was necessary to subtract the mean height before the initial levelling, as the instruments used recorded the minimum height as zero, whereas the segmentation algorithm assumes that the mean height is zero. Therefore, the segmentation would likely fail if the mean height was not subtracted.

3.3.3 Gradient thresholding

To apply the gradient threshold, the gradient map is first calculated using the Sobel gradient. The Sobel gradient is widely used in image processing (Gonzalez and Woods 2008) and uses two 3×3 kernels, shown in figure 3.4 to calculate the x and y gradients. These can then be combined to calculate the gradient magnitude. The gradient is reported as a ratio of height/spacing and is therefore unitless provided the lateral and vertical units are the same (e.g. µm/µm).

The gradient thresholding algorithm then works similarly to height thresholding, except on the gradient map. In this case points above the threshold where marked as features (1) and points below as background (0). The post-processing steps are particularly important for gradient thresholding as the base of the feature is often flat and is likely to be misidentified.

3.3.4 Morphological segmentation

To perform morphological segmentation, MountainsMap software (*Mountains Map 7*) was used. This software provides a fast and stable merging implementation, which is difficult to implement and achieve practical running speeds for in Matlab. MoutainsMap allows merging based on both height and area of regions, by setting thresholds were watershed regions with height or area below the threshold are merged, and can easily perform watersheds on the gradient. Information about the exact merging algorithm used is not available. However, given the option to use both height and area thresholds,

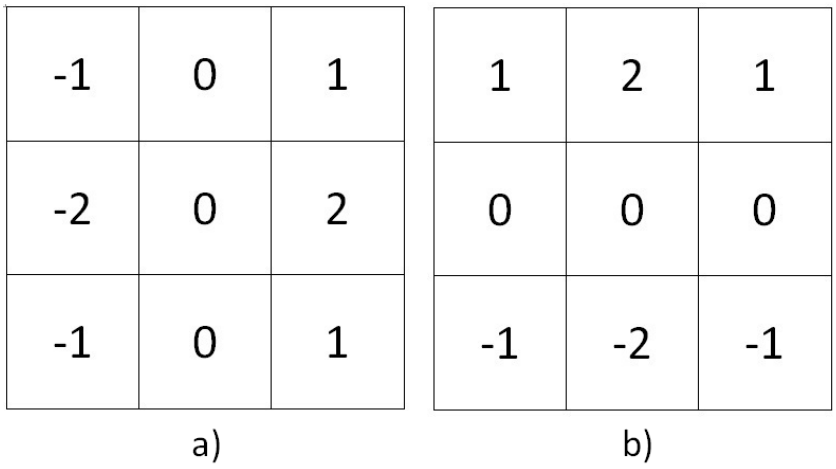

Figure 3.4: 3×3 *Sobel filters. a) x gradient. b) y gradient. Gradient magnitude can be calculated from combination of x and y gradients.*

and the stability of the method, it is assumed a version of Barre and Lopez's (Barre and Lopez 2000) algorithm is used. The diameter of the segmented regions can be determined directly from the software. Alternatively, maps of the segmentation can be exported and analysed further in Matlab.

The software reports the thresholds used as both absolute values and as percentages. Absolute values are preferred in this thesis, particularly for height/gradient thresholds where the result is reported as a percentage of Sz, which is known to be sensitive to noise. Additionally, when segmenting based on the gradient map the height (gradient) threshold is still reported in micrometres. This is obviously incorrect for a gradient and in this thesis the value is reported as a unitless gradient, although the exact definition of the gradient in this case is unknown.

3.3.5 Active contour

The active contours without edges algorithm (Chan and Vese 2001) was used and was implemented using Matlab's built in algorithm. The active contour approach requires an initial contour to be set and a simple square 300 pixels from the image edge was used. This is not the optimal solution and faster convergence could be achieved using a more advanced technique to choose the initial boundary. The algorithm was allowed to run for a maximum of 5000 iterations for initial levelling, which should help ensure convergence. The number of iterations is particularly important for higher smoothing factors which

where found to need longer to converge. After the initial levelling, the segmentation from the initial run could be used to define a contour to re-initialise the algorithm for the final segmentation. This approach provides a more accurate initial contour and so only 2000 iterations were needed for subsequent runs.

3.3.6 Post-processing

The post-processing step aims to correct any misidentified regions and produce a final segmentation map that accurately represents the feature measured. Two steps are used in the post-processing. First, background regions, which are surrounded by feature regions, are filled in and marked as features. This step deals with parts of the feature which were not correctly marked and is particularly important for the thresholding methods, where some regions of the base of the feature are often misidentified.

Additionally, small feature regions are removed and marked as background. This step accounts for the effect of background roughness or other effects that can cause background points to be misidentified. In cases where partial features are present on the boundary, removal of small regions should also remove partial features, either due to their small size or as an additional constraint. The size of region that should be removed is somewhat dependant on the feature size. For most situations, regions below 5000 pixels were removed. In practise, it was helpful to perform filling before removing small regions as filling the main feature helped ensure that it was large enough to not be removed. However, small regions were not removed when levelling, only for the final segmentation. As noted above, removing small regions also removes partial features on the boundary. These can significantly skew the levelling and so should be excluded.

Figure 3.5 shows an example of the combined segmentation and post-processing stages for gradient thresholding. The process is similar for other methods.

3.3.7 Determination of geometric properties

Three geometric properties are considered in this thesis, but there are a wide range of others that may be of interest. Figure 3.6 shows these three properties and gives a graphical indication of how they are calculated. Due to the nominally circular shape of the dimples investigated, it is natural to fit a circle to the feature boundary. In particular, the NPL LSGE toolbox was used to perform this circle fitting (*NPL LSGE library*). From this fitted circle the radius, or equivalently diameter, is defined.

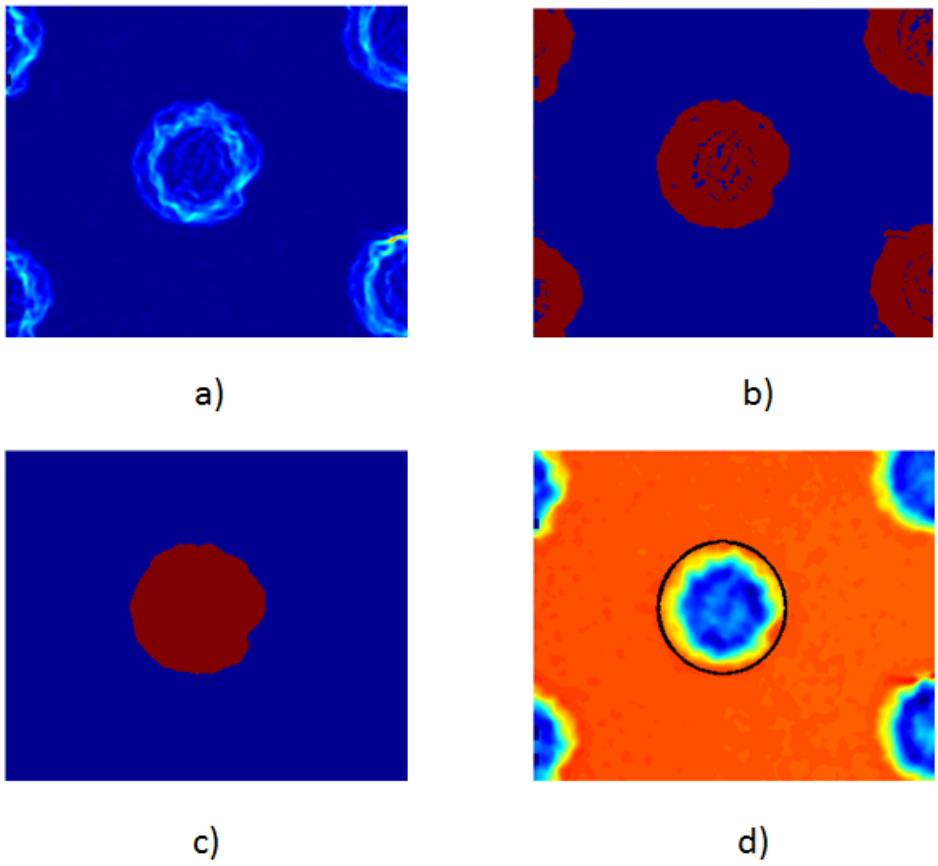

Figure 3.5: Overview of the height thresholding process showing: a) Gradient map of a dimple. b) Initial segmentation after gradient thresholding. c) Final segmentation after removal of small feature regions and filling of enclosed background regions. d) Circle fitted to feature boundary imposed on original topography.

Dimple depth is considered using an areal step-height analogy. The depth is defined as the mean height, assuming the background is levelled to zero, of a circle concentric with the circle fitted to the boundary, but with half the radius. Using a circle of half the radius ensures that the feature walls are excluded from the depth calculation. Various other methods of calculating feature depth, such as using the bearing area curve (ISO 2012), could also be used but were not directly considered here.

Thirdly, the out of roundness is defined as the sum of the maximum positive and negative radial deviations from the fitted circle. The radial deviations of the boundary pixels can be calculated given the circle centre and radius. The difference between the maximum positive and negative values give the out of roundness. Again other definitions

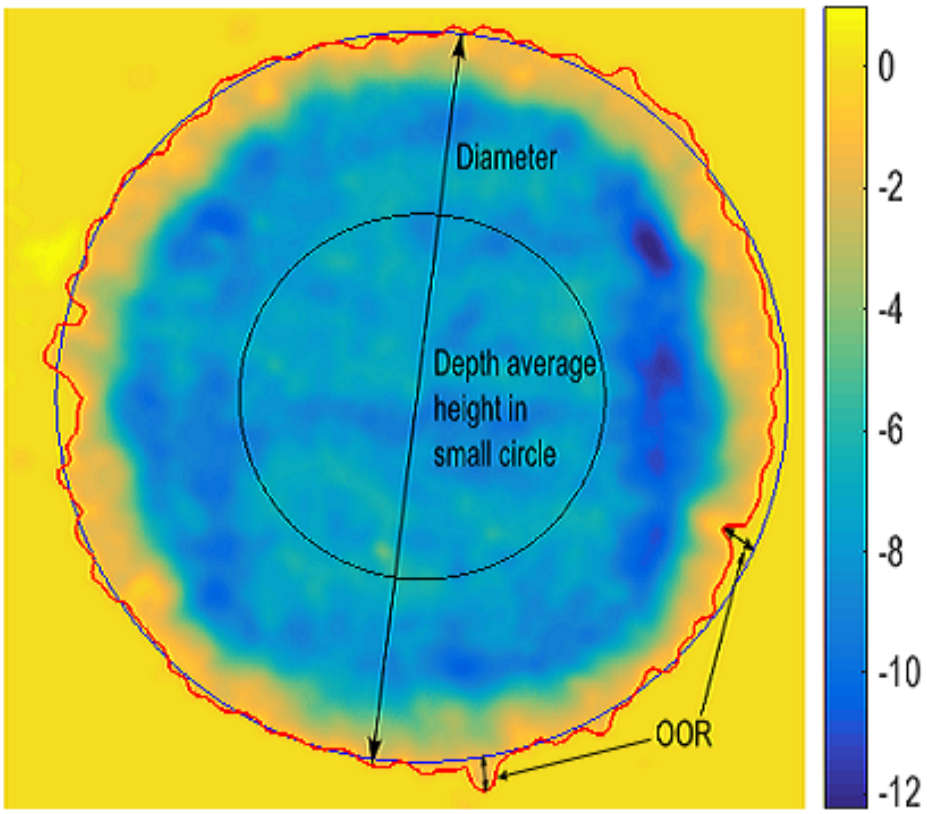

Figure 3.6: Example of geometrical properties of a dimple. Red contour shows extracted feature boundary (height threshold -0.5 μm*). Blue circle is the least squares circle of the boundary. Black contour is the half the diameter and used to calculate depth. Out of roundness (OOR) is calculated as the sum of the maximum radial deviations from the fitted circle.*

of out of roundness could be considered, for example based on maximum inscribed and minimum circumscribed circles. However, given that the radius has been defined in terms of a least squares circle, this definition seems the most natural.

3.4 Summary

This chapter has considered the samples, measurement and analysis techniques which will be used. The main points were as follows:

1. The samples used are laser textured silicon nitride disks consisting of a regular

pattern of pits or dimples, with various different nominal diameters and depths on different disks.

2. Laser confocal and focus variation microscopes are used to measure individual dimples on the surfaces.

3. Characterisation of the geometric properties of the measured dimples, diameter, depth and out of roundness, is performed using four different segmentation methods: height thresholding, gradient thresholding, morphological segmentation and active contours.

Chapter 4

Factors affecting geometric characterisation of structured surfaces

The segmentation methods presented in chapters 2 and 3 are well understood from various fields, such as image processing and computer vision. However, their application to structured surfaces has not been investigated in detail and care is needed to fully understand the physical meaning of the results. In particular, it is important to understand how a feature is defined, as this will have a significant impact on the usefulness of various segmentation methods.

Therefore, this chapter considers the definition of a feature in more detail, before comparing the various segmentation methods and investigating how they respond to changes in their control parameters and the size of filter used.

4.1 Definition of a feature

Before considering the effect of segmentation, it is important to define what segmentation is trying to achieve. Segmentation aims to identify the position, size and shape of the feature. One way to determine whether this is achieved would be to compare the boundary of the segmented feature with the true feature boundary, with any differences indicating an error. In practise, such a comparison is not possible without first considering how the feature is defined.

Holistically, a feature can be defined as a region of the surface with similar values of

some property; this property will generally be height but other local properties could be used. Alternatively, a feature can be defined by its boundary, where the boundary defines some transition in a property of the surface. These definitions can be viewed as equivalent; if there is no transition at the boundary, then regions outside the feature are still similar to regions inside. The problem then becomes one of how to define the transition at the boundary or equivalently how to define what constitutes a region of similar pixels.

One approach, common in computer vision and image processing, is to place the boundary along the ridge with maximum gradient. This approach would be consistent with that used in morphological segmentation and some active contour algorithms, as well as traditional edge detection algorithms from image processing, such as Canny's algorithm Canny 1986. Alternatively, regions of similar pixels could be defined, for example, by thresholding methods or the active contours without edges algorithm Chan and Vese 2001. However, in practise, these definitions do not always agree. In particular, differences often occur at diffuse (non-sharp) boundaries. Such boundaries consist of edges with a significant spatial extent. Depending on how the feature boundary is defined, the boundary can lie at a different position within this diffuse edge.

It is, therefore, important to consider how changing segmentation method and parameters affects the segmentation. If different segmentation choices are equally valid, it could be interpreted as an uncertainty in the geometrical properties of the feature due to uncertainty in the correct specification of the segmentation. However, this viewpoint is not used in this thesis, as there is no easy way to determine if a particular segmentation is valid. Instead, the choice of segmentation method and parameters is viewed as defining the feature, such that changing the method is assessing a 'different' feature. In this case, it is still important to consider the effect of changing the method, as a suitable choice of method should provide the most relevant results.

4.2 Effect of changing control parameters

The effect of changing the control parameters of the segmentation algorithm was investigated using an example dimple of nominal diameter 300 μm and depth 10 μm. Four segmentation methods were investigated: height thresholding, gradient thresholding, morphological segmentation and active contours. For each method, a range of segmentation parameters were investigated. A circle was fitted to the segmentation boundary and the radius and depth of the feature were determined, as detailed in section 3.3. The following sections present these results for the four different methods used.

4.2.1 Height thresholding

Height thresholding was used to segment the surface using a height threshold ranging between -8 and 2 µm. This range was used as initial investigations found that outside this range the segmentation was unsuccessful. Figure 4.1 shows the segmentation boundaries produced by height thresholding at some example threshold values and how they relate to the visible boundaries of the actual dimple. At mid-ranged threshold values, the segmentation boundaries lie somewhere on the sloped walls of the dimple; at very low thresholds they travel into the base of the feature and the dimple boundary is lost. Similarly, at high threshold values, the segmentation boundary begins to expand towards the regions surrounding the dimple.

Figure 4.2 shows a plot of the radius, computed using the circle fitted to the feature boundary as described in chapter 3, against threshold used. The regions, where segmentation degenerates and the dimple radius is lost, are clearly visible on both the left and right-hand sides of the plot. In the middle region, where segmentation is successful, the dimple radius changes gradually with threshold due to the slope of the dimple walls. Any of these intermediate threshold values could provide a valid dimple boundary, which indicates the possibility of customising the segmentation via the threshold value in order to fulfil specific, application dependent characterisation requirements. For the test case, a radius change of about 20 µm is observed in the region corresponding to valid threshold values. It should be noted that the sensitivity of the algorithm is strongly dependent on the geometry (e.g. slope) of the dimple walls, and different topographies may not give the exact same sensitivity to the threshold value.

Interestingly, in figure 4.2 there is also an unstable region at low threshold, before the segmentation degenerates completely. This instability is caused by the boundary the boundary entering the base of the dimple (see the red contour in figure 4.1), where the roughness of the feature base causes the feature boundary to become significantly deformed and unstable. The measured radius can, therefore, increase as well as decrease due to incisions into the centre of the dimple that tend to have a longer length and so have a greater weight when fitting. At lower thresholds, these incisions are then removed as some of the surrounding regions are removed from the feature, conversely resulting in a larger fitted radius. However, it should be noted that the trend is still towards a decreasing radius at lower threshold.

One way of evaluating the quality of the result is to use the standard deviation of the residual to the fitted circle. Figure 4.4 shows the standard deviation of the residual

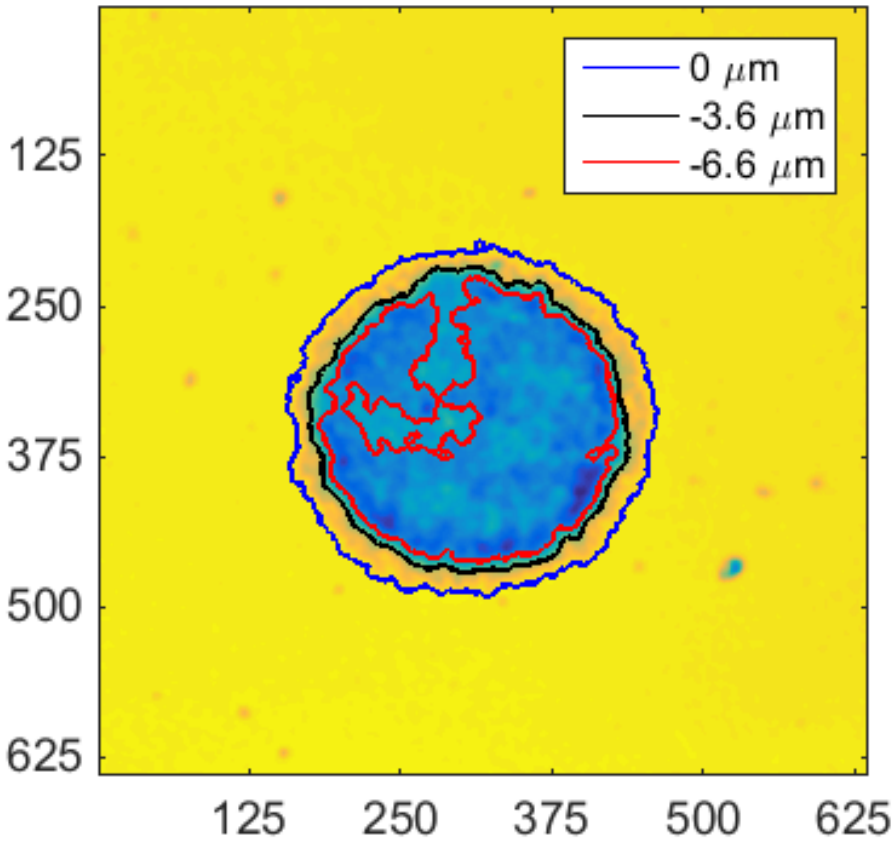

Figure 4.1: Boundaries produced by height thresholding using different threshold values. xy axis in micrometers µm.

as the height threshold changes. There is a significant increase in the standard deviation in the region where the segmentation fails. On the other hand, there is relatively little change in standard deviation where the algorithm is performing properly. It should be noted that this result does not give any indication of the accuracy of the segmentation, just how close to a circle it is. However, sharp changes in the standard deviation still give a useful indication that a significant change has occurred, which may not be immediately noticeable in the measured radius.

The effect of choosing the threshold value algorithmically via k-means clustering ($k = 2$) of image heights is shown by the dashed lines in figure 4.2 and by the black contour in 4.1. The threshold value determined by k-means clustering ($k = 2$) is influenced by the relative number of data points (pixels) being assigned to the two clusters. This may not necessarily correspond to any functionally relevant height in the dimple

topography. To keep the advantages of using an algorithmic approach for determining the threshold while reducing this effect, a k-means clustering with $k > 2$ (see e.g. Senin et al. 2014) could be chosen, as long as a sufficiently robust procedure can be set up under post-processing to merge the excess clusters into a meaningful binary partitioning that correctly discriminates feature and background. However, this choice is not considered and the question of how many clusters to used is not trivial.

Figure 4.3 shows the measured depth of the feature against threshold used. The data follows a similar pattern to that of the radius measurements. This similarity is unsurprising due to the way the depth is defined; the area over which the depth is calculated is dependent on the radius. For the test case, there is less change in depth across the range of suitable threshold values than for the radius measurement. This is because the region at the dimple bottom used to compute the depth is at roughly uniform height and so a small change in the area considered will have minimal effect on the measured depth.

4.2.2 Gradient thresholding

Gradient thresholding was performed using thresholds over a range of gradients between 0.02 and 2. Figure 4.5 shows how the segmentation boundary evolves in gradient thresholding. Both low and high thresholds lead to failure in the segmentation. Thresholds that are too low will lead to many irrelevant topography regions being picked up as high-sloped, even if located in reasonably flat areas outside the dimple. Whereas, Thresholds that are too high will cause the algorithm to miss most of the relevant slopes. Post-processing can mitigate these effects to some extent, but the range of useful threshold values for computing dimple radius is still limited. This behaviour is demonstrated by the change in dimple radius as a function of threshold value, as illustrated in figure 4.6.

As for height thresholding (figure 4.2), the plot in figure 4.6 suggests that, in the range of valid threshold values, gradient thresholding may be adjusted to produce range of radii that depends on which points are considered part of the feature. Algorithmic determination of optimal thresholds may also be attempted, following the same conceptual lines suggested for height thresholding. Again k-means clustering (of gradient values, with $k = 2$) is applicable, as shown by the dashed lines in figure 4.6 and the red boundary at 0.72 threshold in figure 4.5.

Figure 4.8 shows the standard deviation of the residual as the threshold changes. Most noticeably, there is a large increase in the standard deviation at threshold 1.3. This point corresponds with a small decrease in measured radius. Further examination

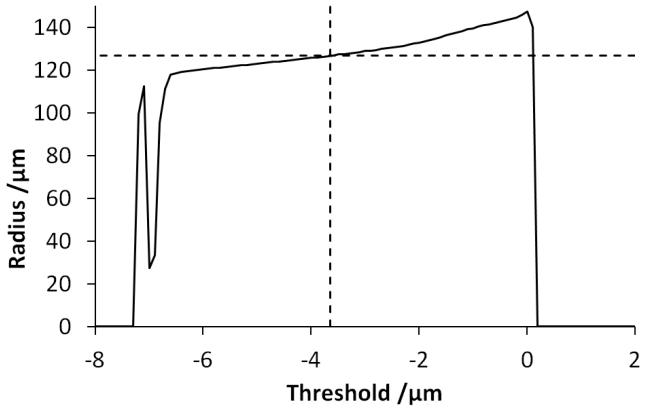

Figure 4.2: Dimple radius as a function of height threshold. Dotted line shows the values that would be achieved by k-means clustering with $k = 2$.

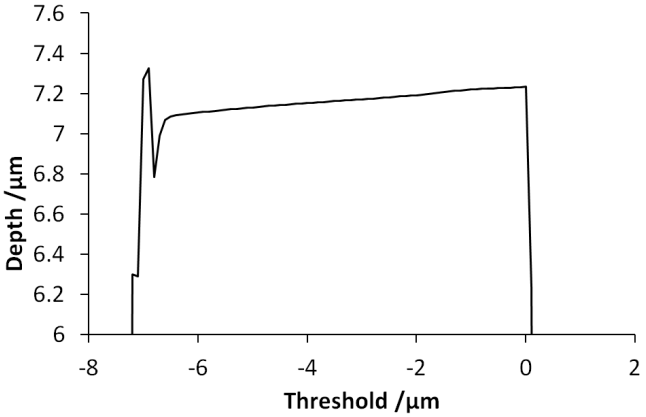

Figure 4.3: Measured depths as a function of height threshold.

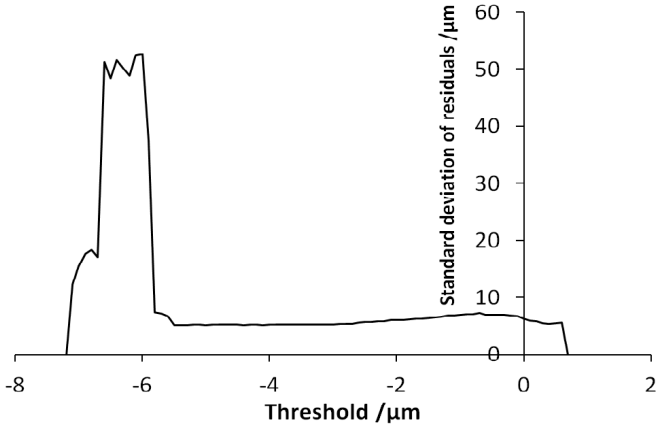

Figure 4.4: Standard deviation of residual as a function of threshold.

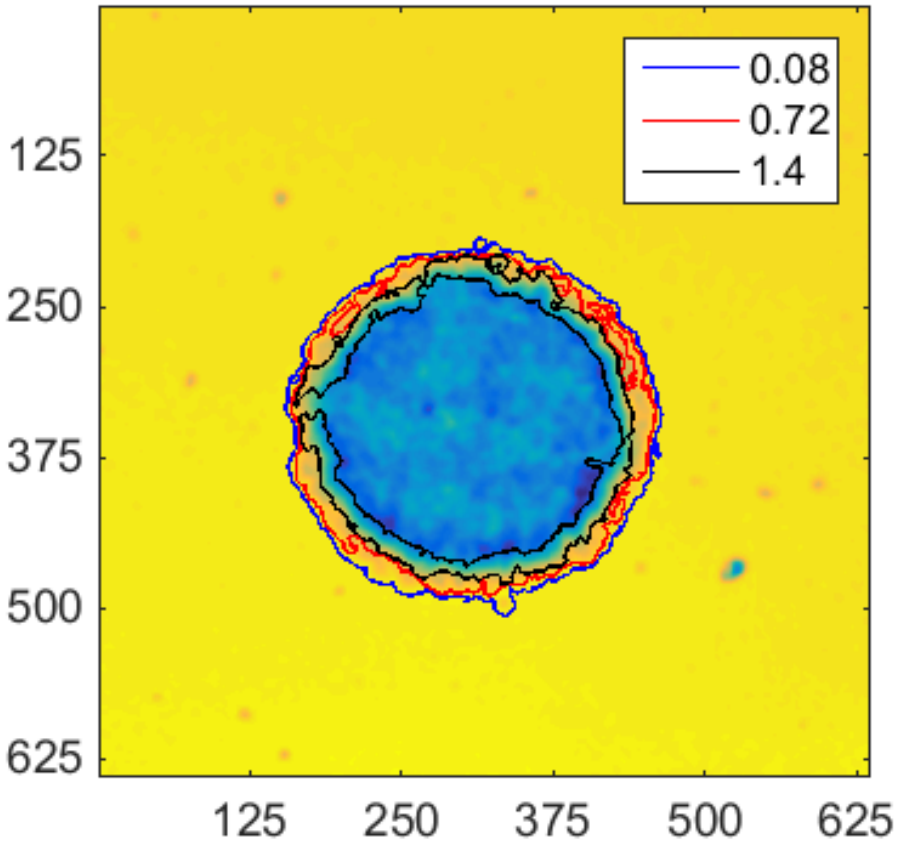

Figure 4.5: Boundaries produced by gradient thresholding using different threshold values. xy axis in µm.

shows that these changes are caused by the increased threshold producing a gap in the detected boundary, resulting in a horseshoe shaped contour shown by the black contour in figure 4.5, for threshold 1.4. This demonstrates that a significant change in the detected boundary, which is easily noticeable in the standard deviation of the residual, can have a relatively small effect on the diameter, which could easily be missed. Additionally, there is a significant increase in standard deviation at about threshold 0.6. While this value does not correspond to any jump in diameter or significant change of measured boundary, it can be seen from figure 4.5 that boundaries below this value, such as the blue contour, are much smoother than those with higher thresholds, such as the red contour. This effect is likely because, as the threshold increases, more and more regions with low gradient are removed from the feature. It is not clear why there is a jump at threshold 0.6. One possibility is that on this dimple a particularly large region is removed at this threshold.

The large incision at the top right of the red contour in figure 4.5 is a good candidate for this effect.

Figure 4.7 shows the dependence of the depth on threshold value. This plot follows the same pattern as for height thresholding, with the depth giving a near constant value over the range where the threshold is generating a valid segmentation result. A larger step in depth values corresponds again to where the maximum of circle fitting residuals is located.

4.2.3 Morphological segmentation

Due to how morphological segmentation, and in particular watershed merging, are defined, changing the threshold parameters can only cause watershed lines, i.e. region boundaries, to be added or removed, not distorted. Therefore, provided that the segmentation performs correctly, it will not change with threshold. Morphological segmentation on the image gradient with an area threshold on minimum watershed size of 5 % was found to effectively segment the surface. Additionally, a gradient threshold on the watershed boundary could be used, with any threshold between 0 and 1.8 having no effect on the result. For larger gradient thresholds, no feature was detected. Similarly, a range of area thresholds could achieve the same effect. It was found that above 11 % area threshold no segmentation could be achieved. Also, for area thresholds below 2.5 % the feature was not fully identified and consisted of multiple regions. In some cases, additional post-processing could be used to merge these regions and alternatively, a combination of area and gradient thresholds was sometimes effective.

Once the feature was successfully identified, it was measured as having a radius of 124.3 µm and a depth of 7.19 µm.

4.2.4 Active contours

The active contour without edges algorithm was used to segment the surface with smoothing coefficients, μ in equation 2.5, ranging between 0 and 30. Figure 4.9 shows a plot of the dimple radius as a function of the contour smoothing factor. The radius decreases slightly with increasing smoothing factor within the range of definition of the control parameter; this decrease is to be expected because high smoothing factors induce the formation of shorter contours, which will be fitted to circles with smaller radii. It is suspected that the observed variations in the radius are due to protruding and intruding deviations in the contour being smoothed out. Protruding deviations will be preferentially removed by the

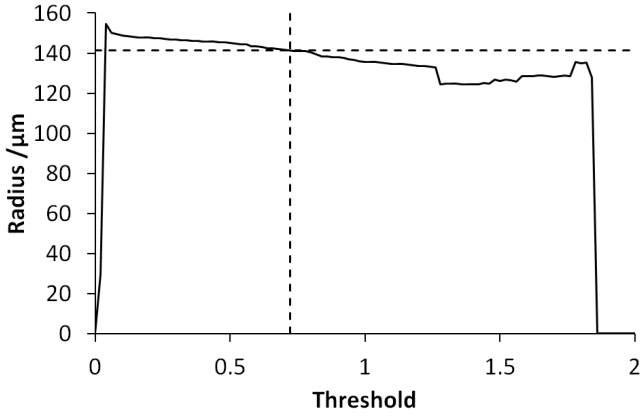

Figure 4.6: Measured radius as a function of gradient threshold. The dotted line indicates the values given by k-means clustering with $k = 2$.

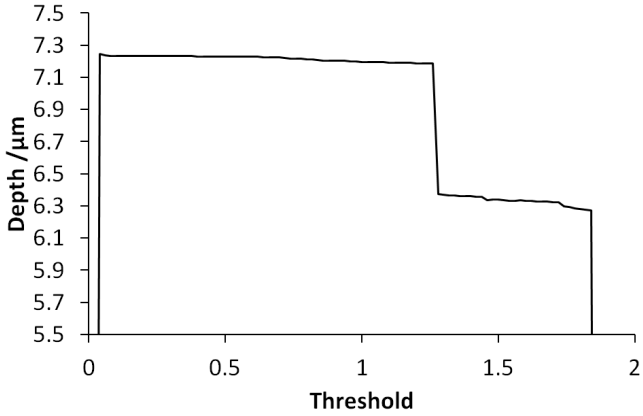

Figure 4.7: Measured depth as a function of gradient threshold.

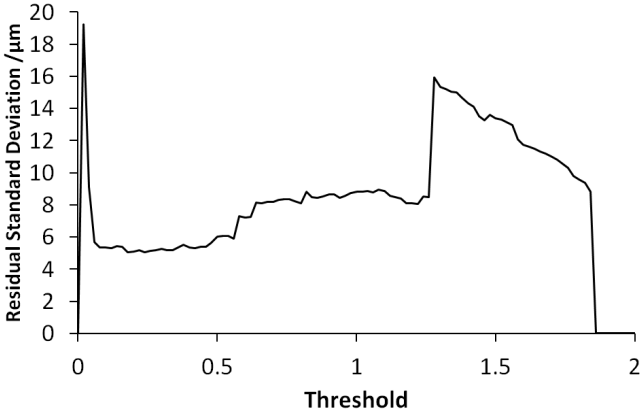

Figure 4.8: Standard deviation of residual as a function of threshold.

algorithm, as they tend to increase contour length more than intrusions, resulting in the general trend of decreasing radius.

It was found that the number of iterations required for convergence increased with smoothing factor. Therefore, the maximum smoothing factor was chosen in order to keep the number of iterations required for convergence within feasible limits. Within the chosen range of the smoothing factor, only valid binary segmentations are generated.

Figure 4.10 shows dimple depth as a function of smoothing factor using active contours. As can be seen, depth is essentially constant. This is expected given the very small change in measured radius, which implies a small change in the fitted circle.

Figure 4.11 shows the standard deviation of the residual against smoothing factor. The general trend is that the residual decreases as smoothing factor increases, which is as expected as deviations are smoothed out by the higher smoothing factor. There is a large spike at $\mu = 22$, which may be an indication that the algorithm has failed to converge at this point. However, this effect is not noticed at higher values and appears to have had minimal influence on radius and depth.

In the future, it may be of interest to investigate how changing the initial contour affects the final segmentation and the time for the algorithm to converge. However, for the current tests, the same initial contour was used for all runs, consisting of a square towards the edge of the image. This contour was chosen as initial tests found that, the initial contour position had no effect on the final segmentation and only effected convergence time, which is not a primary concern for this analysis.

4.3 Comparison between segmentation methods

Comparison between the different segmentation methods is not straightforward. For example, it is not possible to make a direct comparison between dimple radii or depths obtained with each segmentation algorithm, at any given value of the control parameter, because the control parameters themselves are not directly comparable between the different segmentation methods. However, an overall assessment of the algorithms' behaviour can be obtained by considering the distributions of the radius and depth values generated by each segmentation algorithm over the range of control parameters used. Box plots of the dimple radii for each segmentation method are shown in figure 4.12; depths are shown in figure 4.13. Each plot provides an indication of range and quartile distributions. Wider distributions are an indication of higher sensitivity of the algorithm to the main control parameter. High or low sensitivity is not necessarily an advantage

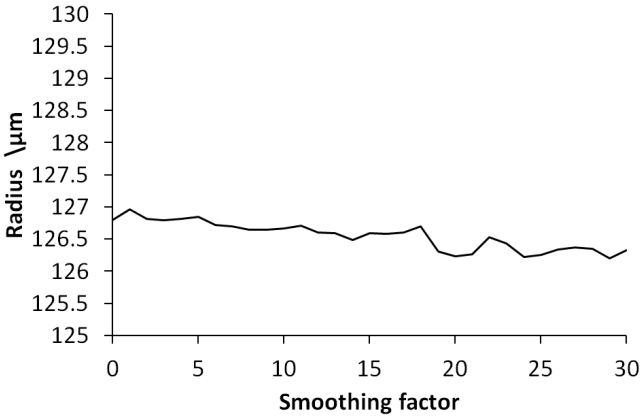

Figure 4.9: Measured radius as a function of smoothing factor.

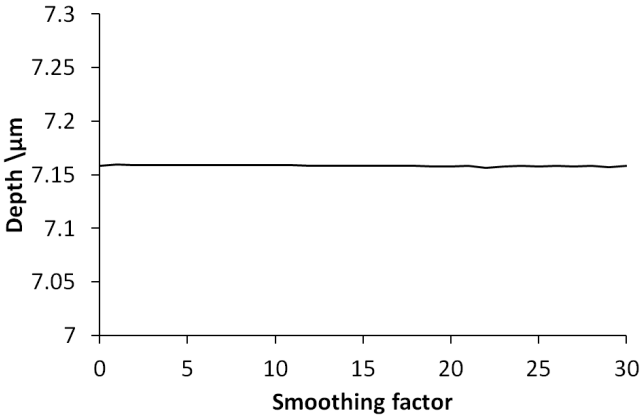

Figure 4.10: Measured depth as a function of smoothing factor.

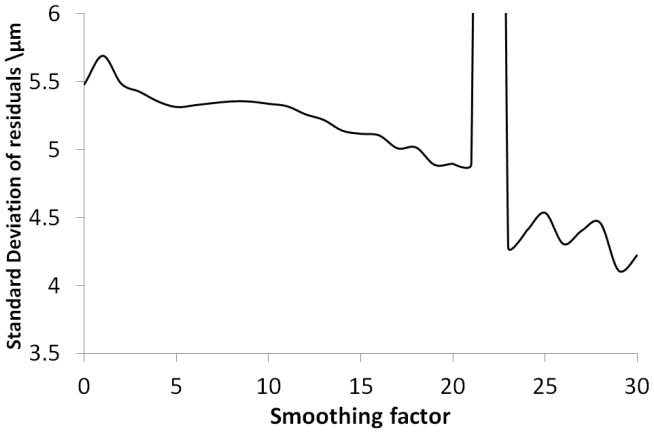

Figure 4.11: Standard deviation of residual as a function of smoothing factor.

or disadvantage. On one hand, higher sensitivity gives greater ability to adjust the segmentation method to provide a relevant result. However, increased sensitivity also makes choosing an appropriate set-up more challenging.

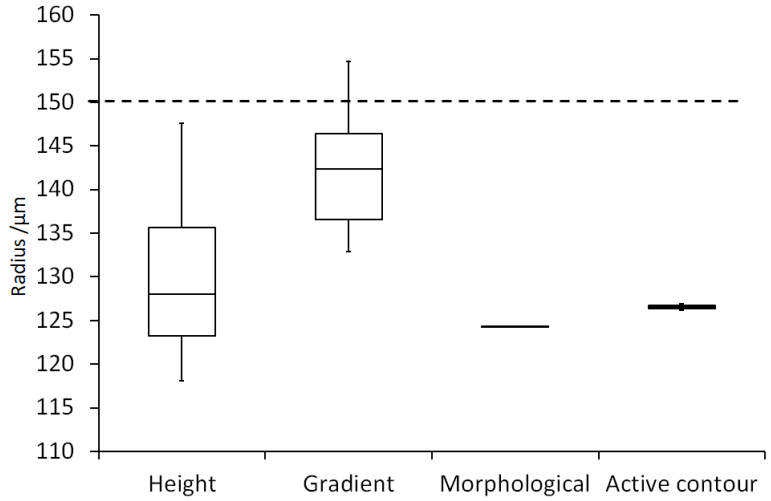

Figure 4.12: Box plot showing the range of achievable radius measurements for each method. The dotted line shows the nominal radius of the dimple.

For the dimple radius, height and gradient thresholding show the highest sensitivity, while morphological segmentation and active contours have lower sensitivity, indicated by the more constant outputs. These results may suggest that morphological segmentation and active contours are easier to set-up. However, height and gradient thresholding, while possibly being more difficult to adjust and optimise, provide greater flexibility and thus greater control over the physical diameter measured. For the test case, both active contours and morphological segmentation locate the feature boundary near the centre of the dimple side surfaces; this boundary corresponds, approximately, with standard definitions of an edge in image processing. However, many mathematical models that are currently applied to model friction reduction on structured surfaces Etsion et al. 1999; Etsion 2013 consider the dimple radius as the top of the dimple, where it meets the background surface. Indeed, height and gradient thresholding could be set-up to provide a corresponding radius, by choosing an appropriate control parameter value. Obtaining equivalent behaviour with active contours or morphological segmentation would be more difficult.

For the depth, the results follow a similar pattern to the radius measurements, although the depths are generally more similar than the radius measurements. This equivalence is to be expected given that depth computation is less sensitive to variation in the

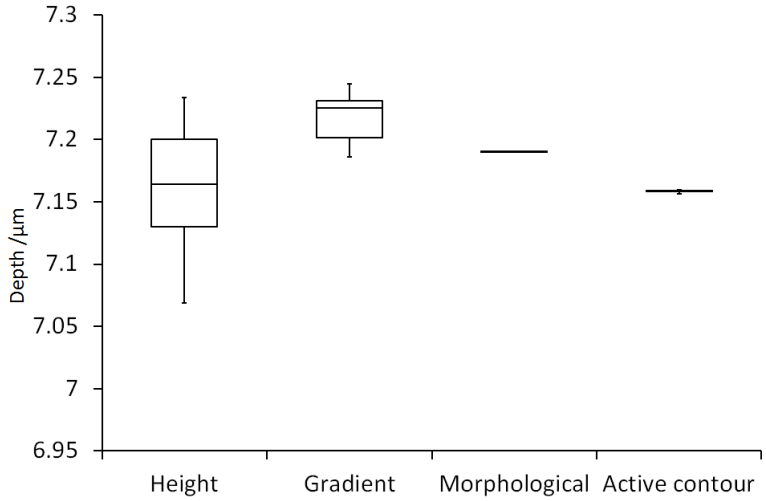

Figure 4.13: Box plot showing range of achievable depth measurements for each method. These heights are all significantly lower than the nominal depth of 10 µm

control parameters, as shown in section 4.2.

4.4 Effect of filtering

Beyond the control parameters of the segmentation method, other choices in the feature identification algorithm can have an effect on the final geometric properties of the feature. One of the most obvious choices is the choice of filter. As mentioned in section 2.4.2, filtering the surface is necessary to remove noise and high spatial frequency components that cannot be measured accurately by the measurement instrument. However, Gaussian filters which are commonly used are known to distort the position of edges and so their effect on feature size should be considered.

In order to estimate the effect of filtering on the surface, the analysis of section 4.2 was run again using a range of different cut-offs for the Gaussian filter. The resulting diameters from this analysis are shown in figures 4.14 to 4.17.

For the height threshold, there is little change to the maximum and minimum usable thresholds. However, there is some change in measured diameter with filter size. It is interesting to note that this change is not consistent with changing threshold height. The maximum change in diameter across the filters investigated here is about 4 µm. Although, in most cases it is approximately 1 µm. Additionally, there is not always a linear relationship between filter size and diameter; sometimes a smaller filter can result

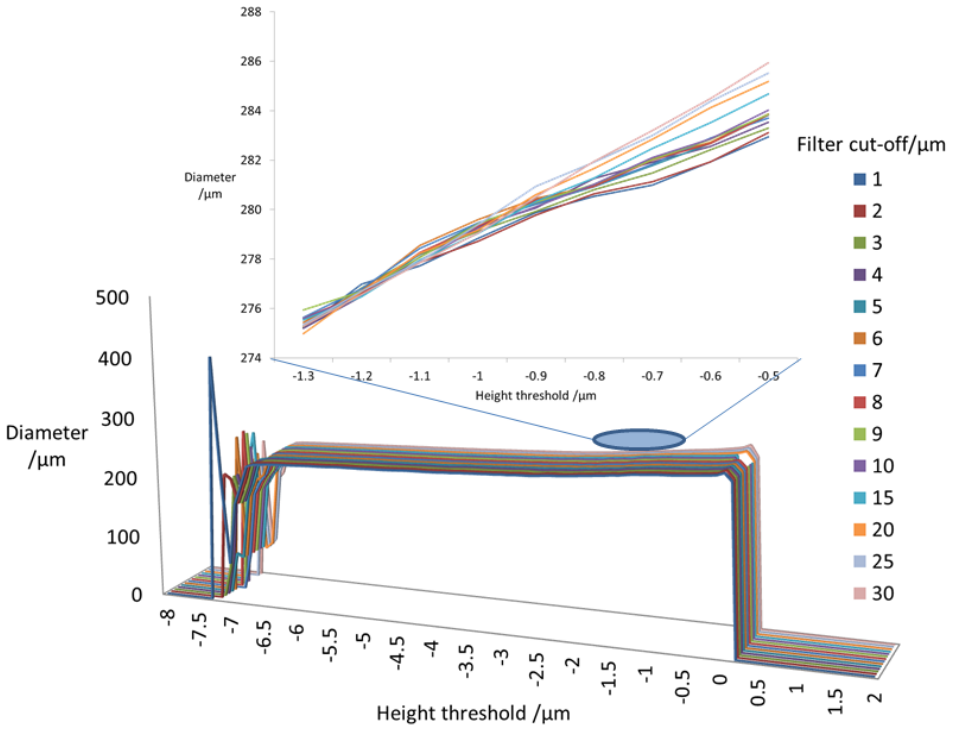

Figure 4.14: Effect of height threshold and filter cut-off on measured diameter

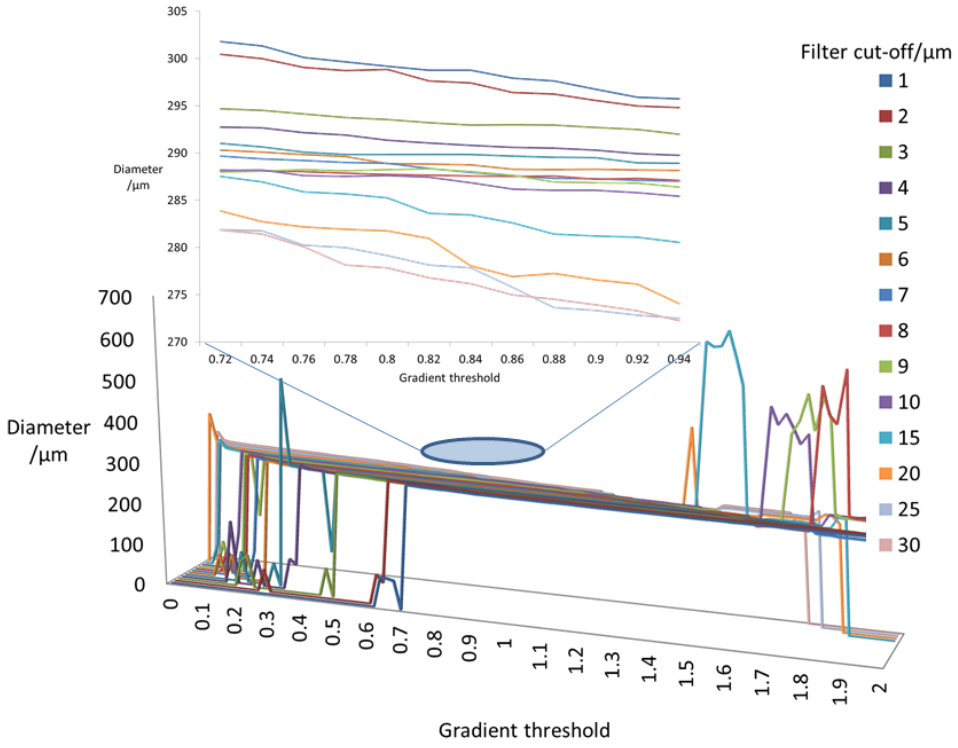

Figure 4.15: Effect of gradient threshold and filter cut-off on measured diameter

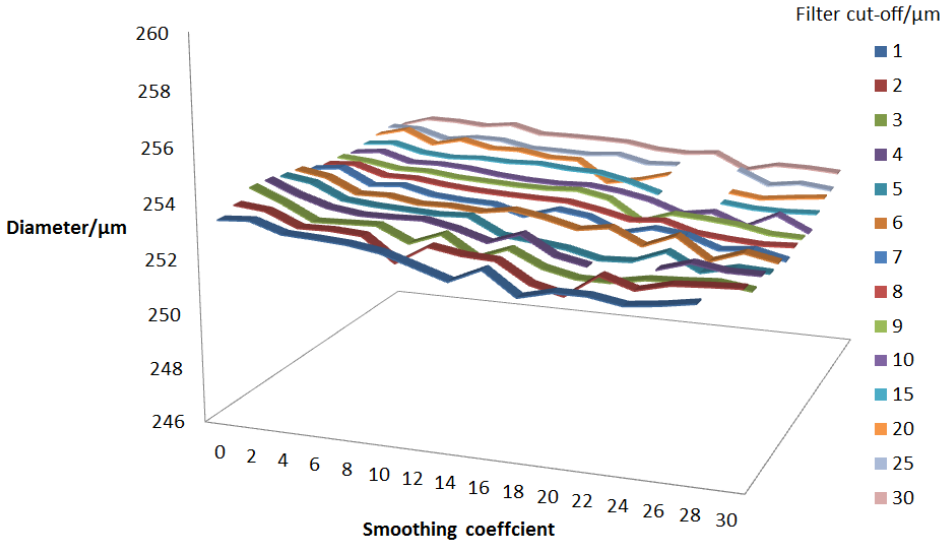

Figure 4.16: Effect of active contour smoothing coefficient and filter cut-off on measured diameter. Some points where the algorithm failed to converge have been removed.

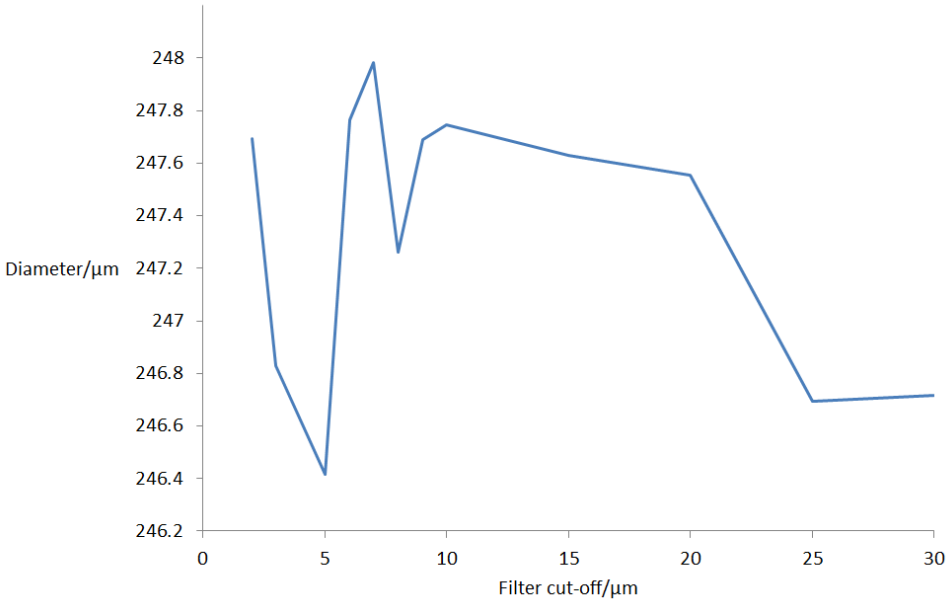

Figure 4.17: Effect of filter cut-off on measured diameter using morphological segmentation.

in a larger measured diameter. These effects are believed to be due to the interaction of local topography with the filter. Near the top of the feature where the gradient begins to decrease, there will be higher curvature and the filter will have a greater effect. Additionally, the lower slope near the top of the feature mean that small changes in height will result in greater changes in diameter. Similarly, changing filter size can move local undulations in or out of the filter and so can cause the surface to be distorted in many different ways.

For the gradient threshold, the change in diameter with gradient is more severe, approximately 20 μm. Additionally, the range of gradients where a successful segmentation can be achieved changes significantly. For small filters, local undulations cause high gradients and the segmentation can fail at low thresholds, as a significant amount of the background is detected as feature. These high local gradients are smoothed out as filter size increases and a lower threshold can be used to get a successful segmentation. However, at high thresholds the segmentation may fail earlier as the high gradients are smoothed out.

For the active contour method, there is little change in measured diameter with filter cut-off, with a change of less than 1 μm as the filter size changes. From the results using the height threshold, it is known that the surface topography must change more than this amount. Therefore, it can be concluded that the height of the threshold position using the active contour changes with filter size. What is interesting is not so much that this effect occurs, but that it occurs in such a way as to keep the diameter approximately constant.

The effect of filtering on morphological segmentation is shown in figure 4.17. As with the other methods, the filter changes the position of the ridge of maximum gradient. Therefore, the position of the watershed lines and measured diameter changes with filter size, varying in a range of approximately 1 μm. As before, changing the threshold values for the merging algorithm does not effect the boundary position, provided that the segmentation is effective. However, the range of values that will achieve a successful segmentation will change with filter size.

4.5 Conclusions

This chapter has considered the effects of changes in the algorithm control parameters and filter size on the geometric properties of a structured surface. Based on this analysis it can be concluded that:

1. The choice of segmentation method strongly effects the geometric properties, both in terms of absolute value and the effect of changes in the algorithm control parameters. These results were more or less as expected in that morphological segmentation and active contour methods are insensitive to changes in the control parameters, resulting in no and very small changes in the feature boundary respectively. Thresholding methods, on the other hand, are far more sensitive as they gradually add or subtract from the feature as the threshold changes, whereas the other methods place the boundary at more or less fixed positions.

   However, these results are somewhat dependant on the feature being investigated. Square or triangular features are likely to be more affected by the smoothing factor in the active contour method than circular features, which have no sharp corners. Similarly, the sensitivity of the thresholding methods is closely related to the slope of the respective map. Features with more vertical walls will be less influenced by the value of the threshold used.

2. It was also shown that, for the thresholding methods particularly, the algorithms can produce a poor or incorrect segmentation while still providing an answer. If using these approaches on other surfaces, care should be taken to ensure that thresholds where poor segmentation occurs are avoided. The values where this is the case will depend strongly on the particular surface being investigated, but can often be detected by an increase in the standard deviation of the residual, even when the feature radius does not change significantly.

3. The filter size used was also shown to have a significant effect on the geometric properties. The effect of filter size was highly dependent on the segmentation algorithm used, with a range of approximately 20 μm over the filter sizes investigated using the gradient threshold method compared to approximately 2 μm for the height thresholding and morphological segmentation, and even less for active contours. Additionally, there was no linear or other simple relationship between filter size and diameter. While initially a strong trend had been expected, it appears that that local topography variations are very significant in determining the effect of filtering.

4. It is important to note that there are many other factors, such as pre-processing steps, levelling and circle fitting algorithms, that will also effect the final geometric properties of a feature. The effect of all these factors will vary from surface to surface

and cannot easily be accounted for without significant effort. A simpler alternative is to keep the analysis method as consistent as possible and note that changing the method will lead to a change in the results.

5. In general, the difficulty in choosing suitable control parameters or even the general algorithm when measuring a structured surface could be considered to contribute to the uncertainty in the resulting geometric properties. However, such an approach has not been considered here, and the uncertainty is only defined with respect to a particular characterisation algorithm.

Chapter 5

Assessment of uncertainty of geometric properties of structured surfaces by repeated measurement

It is important to assess the uncertainty associated with any measurement of structured surfaces, as without an associated uncertainty the practical use of a measurement is limited. The assessment of uncertainty for structured surfaces is a challenging problem as the identification of structured surfaces often involves many non-linear operations which are not easily considered using the standard GUM (JCGM 2008a) approach. One possibility to avoid these issues is to estimate reproducibility of the geometric properties being characterised directly, by repeated measurement of the surface. Such reproducibility studies are often used in industry and are relatively easy to perform, but no research using the measurement reproducibility for structured surfaces has been published and care is needed to ensure all sources of reproducibility are assessed.

Measurement reproducibility can be estimated by the standard deviation of repeated measurements under slightly varying measurement conditions, where all conditions not varied are kept the same (JCGM 2012). For these measurements the main factor that is varied is the position of the feature in the field of view, which simulates making measurements where the position of the feature is not carefully controlled. This is as opposed to assessing the measurement repeatability, which is defined as the uncertainty in repeated measurements where all measurement conditions are maintained (JCGM 2012). While assessing repeatability only may appear to give a lower uncertainty, it does not account for the uncertainty in positioning that is inevitable in any real measurement and, therefore,

may underestimate the uncertainty.

This chapter applies such an approach to structured surfaces, estimating reproducibility in the geometric characteristics of a feature based on repeated measurements. Furthermore, the presence, or lack, of any position dependence on the geometric properties is considered, the effect of varying threshold values on the uncertainty is investigated, as well as the effect of varying segmentation method and measurement instrument.

5.1 Assessment methodology

In many cases, the main contributor to measurement uncertainty is the measurement reproducibility. Providing that tests are carefully planned, measurement reproducibility can account for the main uncertainty contributions in measurements of structured surfaces, by calculating the geometric properties of surface features as they are moved around the field of view.

As a test case, three disks with differing nominal diameters ($50, 150, 300$ µm) but the same nominal depth and area density (10 µm and 20 %) are used. On each disk, three dimples are selected and measured repeatedly using the focus variation microscope. Each dimple was measured on a 5×5 grid of positions with ten repeat measurements at each location. Grid spacing was chosen for each disk, as shown in figure 5.1, to ensure that the dimple travelled across the entire field of view while remaining entirely within it. To ensure any effect from the reproducibility of the positioning system is accounted for, the microscope moves between each measurement, completing one 5×5 grid before starting the next repeat.

The geometric properties of the features are then assessed as described in chapter 3. For this test, pre-processing consists of filtering with a 1.6 µm standard deviation Gaussian filter and levelling against the background only. Segmentation was initially performed by gradient threshold using a threshold on the gradient magnitude of 0.3. Enclosed regions were then filled and small regions removed to produce a final segmentation. For the smaller dimples, multiple dimples where present in the same field of view. The dimple of interest was determined by calculating the expected shift from the centre and taking the dimple nearest that point.

Once the feature was segmented, the diameter and out of roundness was calculated. Other parameters could also be considered using the same approach but only these two are used in this analysis. Once calculated, the diameter was compared to the nominal value to create a diameter error, which allows for easier comparison between disks.

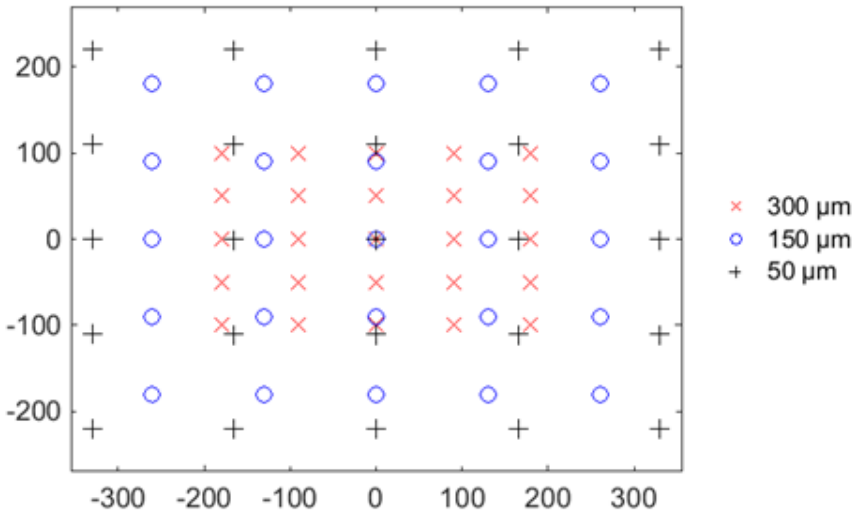

Figure 5.1: Grid of dimple centre positions for different nominal diameters in the field of view, in µm*. Axes cover the entire* 715 µm × 544 µm *field of view of the instrument.*

Reproducibility is assessed in terms of standard deviation associated with each parameter. Additionally, it may be useful to consider whether there is any position dependence across the field of view. Position dependence can be useful to assess whether there are any systematic distortions in the measurement across the field of view, and can be assessed by measuring the mean for each grid position. These means can be used to create a correction map for each dimple. If there is a similar pattern across the three dimples measured on each disk, then an average can be taken to create a distortion map for that disk.

Additionally, it is important to consider how the measurement reproducibility of a single dimple compares with the reproducibility of the manufacturing process. Additional measurements were made to assess the reproducibility in the manufacturing. A random selection of 100 dimples were chosen on each disk and measured once using the same set-up and analysis procedure as above. This sample was then used to estimate the average parameters over the disk and their reproducibility. This manufacturing reproducibility can be compared to the results on a single dimple to determine the significance of the measurement reproducibility. If the measurement reproducibility is small compared to the manufacturing reproducibility, then type A uncertainties are likely to dominate the uncertainty in mean values of the features and measurement of the manufacturing quality not be strongly affected by the measurement reproducibility.

5.2 Assessment of measurement reproducibility

The results of the reproducibility tests on a single dimple are summarised in table 5.1 and figure 5.2. One point to note is that the results for out of roundness tend to have greater standard deviation than for diameter. This increased variability is to be expected, as the out of roundness is sensitive to extreme values. Therefore, small changes in the detected boundary could have a significant effect on the out of roundness, whereas the same changes would be insignificant when calculating the diameter through circle fitting.

Table 5.1: Summary of results of reproducibility tests showing diameter error and out of roundness (OoR)

Diameter/µm	50			150			300		
Dimple no	1	2	3	1	2	3	1	2	3
Mean dia. error/µm	-5.62	-2.08	-4.83	6.29	4.24	6.59	2.45	6.70	-0.16
Std. dia. error/µm	0.17	0.16	0.14	0.14	0.14	0.11	0.083	0.090	0.082
Mean OoR/µm	6.97	8.92	7.67	18.70	11.40	13.51	14.52	10.32	8.49
Std. OoR/µm	0.25	0.25	0.44	0.68	0.23	0.39	0.53	1.23	0.25

In terms of uncertainty, the mean values are of little relevance. However, the standard deviation provides an estimate of the standard uncertainty ($k = 1$) due to the reproducibility of the measurement. There are small differences among the standard uncertainties of each set of three dimples, due to variations in local topography and the finite number of measurements made. However, a single uncertainty estimate for each nominal diameter is highly desirable as otherwise uncertainty would have to be estimated for each individual dimple on the surface, which would be impractical. Therefore, a conservative uncertainty estimate is made by taking the maximum standard deviation of the three dimples at each nominal diameter. Assuming a Gaussian distribution, this approach gives expanded uncertainties ($k = 2$) for the 50, 150 and 300 µm dimples respectively as: ±340 nm, ±280 nm and ±180 nm in diameter and ±880 nm, ±1.4 µm and ±2.5 µm in out of roundness.

To determine whether the measurement reproducibility is significant, it should be compared to the manufacturing reproducibility, i.e. the measurement of multiple different dimples. Table 5.2 and figure 5.3 show the results using the set of 100 different dimples. These results show no obvious pattern in the diameter error between the three disks, but the out of roundness results increase with increasing nominal diameter.

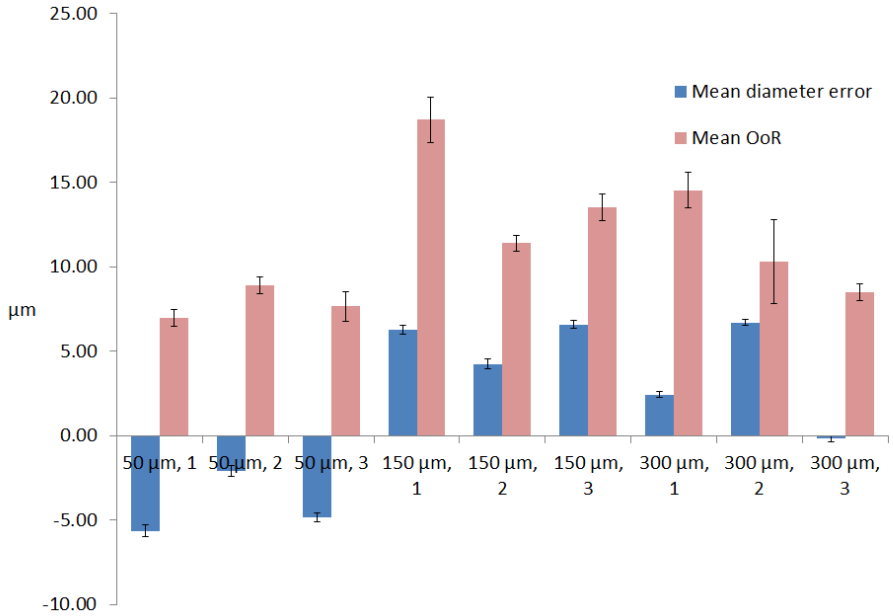

Figure 5.2: Mean diameter and out of roundness (OoR) of measured dimples. Error bars given by 95 % confidence interval.

Table 5.2: Mean and standard deviation (std) of diameter error and out of roundness (OoR) for each set of samples.

Nominal diameter/µm	50	150	300
Mean dia. error/µm	-4.64	5.79	1.67
Std. dia. error/µm	3.38	1.69	4.19
Mean OoR/µm	8.63	15.36	18.29
Std. OoR/µm	3.18	2.59	5.45

In order to assess the significance of measurement uncertainty on a single dimple compared to manufacturing uncertainty of multiple dimples, the variances of the two sets of data are compared. Table 5.3 shows the variance due to measurement reproducibility over the variance due to manufacturing as a percentage, calculated as

$$R = 100\left(\frac{\sigma^2_{meas}}{\sigma^2_{manuf}}\right) \tag{5.1}$$

where R is the ratio and σ^2_{meas} and σ^2_{manuf} are variances due measurement and manufacturing respectively. For the measurement reproducibility, the largest variance of the three dimples at each size was again chosen in order to account for the worst case scenario. For diameter, the effect of measurement reproducibility is small compared to the manufactur-

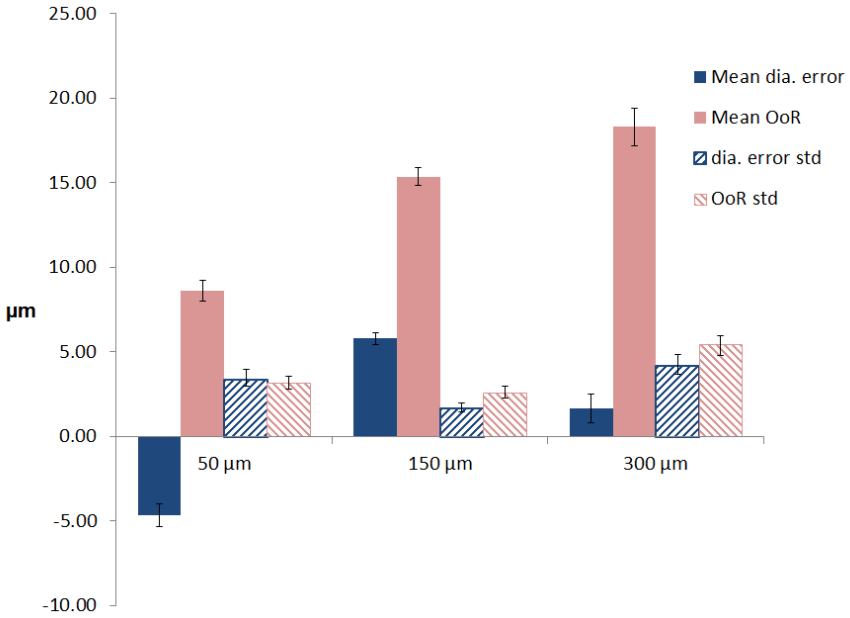

Figure 5.3: Means and standard deviation (std) values for diameter error and out of roundness (OoR) for each set of samples. Error bars indicate 95 % confidence intervals.

ing reproducibility, less than 1 % of the variance of the manufacturing process. For out of roundness, the effect of measurement reproducibility is slightly larger, approximately 5 % of the variance due to manufacturing, although this is still relatively small.

Table 5.3: Ratios of the variance for diameter error and out of roundness (OoR) for the three dimple sizes. The ratio is calculated as the maximum variance due to measurement reproducibility on a single dimple as a percentage of the total variance when measuring multiple dimples.

Diameter/µm	50	150	300
Dia. error ratio/%	0.26	0.71	0.046
OoR ratio/%	1.96	6.83	5.12

5.3 Assessment of position dependence

Position dependence of the measurements is also considered. If this is significant compared to measurement reproducibility then correction for measurement position may be useful. Figures 5.4 to 5.9 show the mean and standard deviation of the diameter using bicubic

interpolation between measurement positions. From these figures, there appears to be a weak pattern in the diameter between dimples of the same diameter. While there are some differences between the three dimples, this is to be expected due to the differences in geometry, exact position of the dimples and reproducibility of the measurement system. The standard deviation does not show any consistent pattern across the dimples. It may be possible to create a distortion map to correct the diameter using this data. However, this correction would be relatively small and, without further information on the cause of the relationship, it is not certain that other features on the surface would also follow the same pattern. Therefore, no correction was performed and the effect of position dependence is considered as part of the uncertainty.

For out of roundness, there is no obvious position dependence in the data. This suggests that the repeatability at a constant position is more significant than any position dependence for out of roundness. Therefore, no correction of the out of roundness is possible with these measurements.

5.4 Comparison of segmentation methods

It is expected that the choice of segmentation method will also effect the measurement reproducibility, as changes in measured heights will have a more or less significant effect depending on segmentation method. To demonstrate this issue, the analysis of the 300 µm dimples measured in section 5.2 was repeated using a height threshold of -0.5 µm, rather than a gradient threshold. The results of the height and gradient threshold approaches are shown in table 5.4, with an expanded uncertainty ($k = 2$) using the height threshold of ± 0.54 µm in diameter and ± 1.9 µm in out of roundness. This process could also be performed with any other segmentation method.

Table 5.4: Comparison of measurement reproducibility using height and gradient thresholds for the 300 µm *dimples.*

Method	Gradient			Height		
Dimple no	1	2	3	1	2	3
Mean dia. error/µm	2.45	6.70	-0.16	-4.45	-1.01	-5.41
Std. dia. error/µm	0.083	0.090	0.082	0.13	0.13	0.21
Mean OoR/µm	14.52	10.32	8.49	13.93	11.40	8.67
Std. OoR/µm	0.53	1.23	0.25	0.33	0.31	0.26

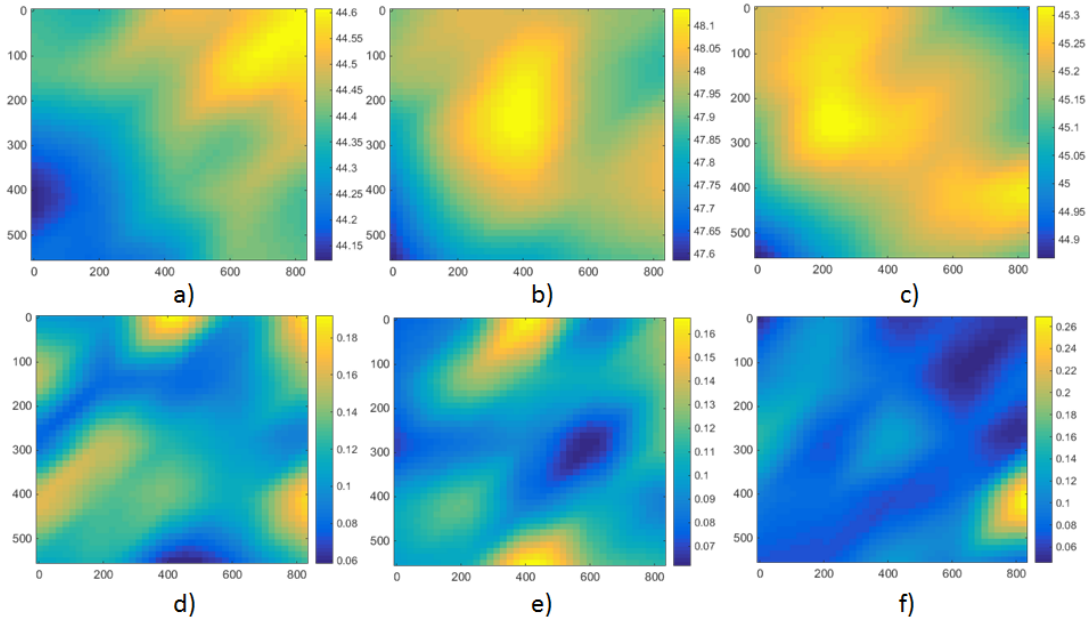

Figure 5.4: Position dependence of diameter for 50 µm *dimples, interpolated using bicubic interpolation. a-c) mean diameter, d-f) standard deviation of diameter. All units in* µm.

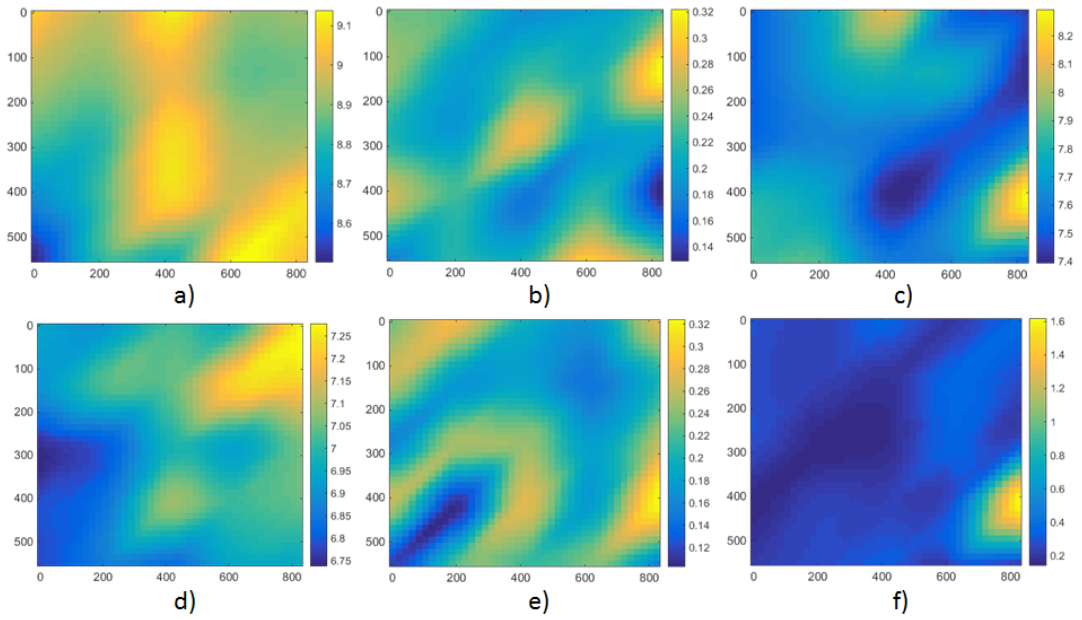

Figure 5.5: Position dependence of out of roundness for 50 µm *dimples, interpolated using bicubic interpolation. a-c) mean out of roundness, d-f) standard deviation of out of roundness. All units in* µm.

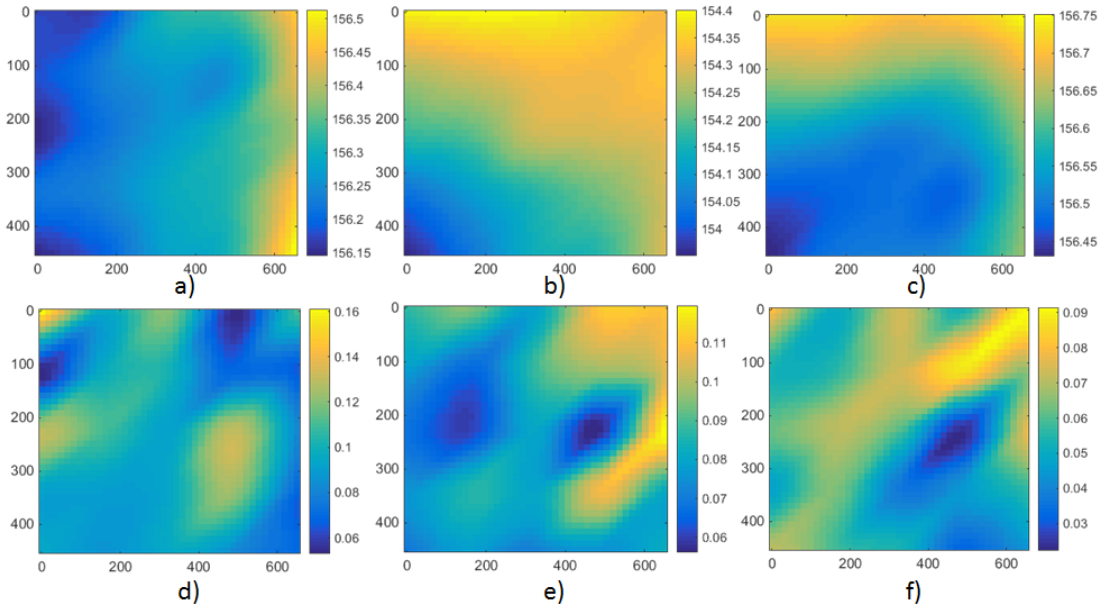

Figure 5.6: Position dependence of diameter for 150 µm *dimples, interpolated using bicubic interpolation. a-c) mean diameter, d-f) standard deviation of diameter. All units in* µm.

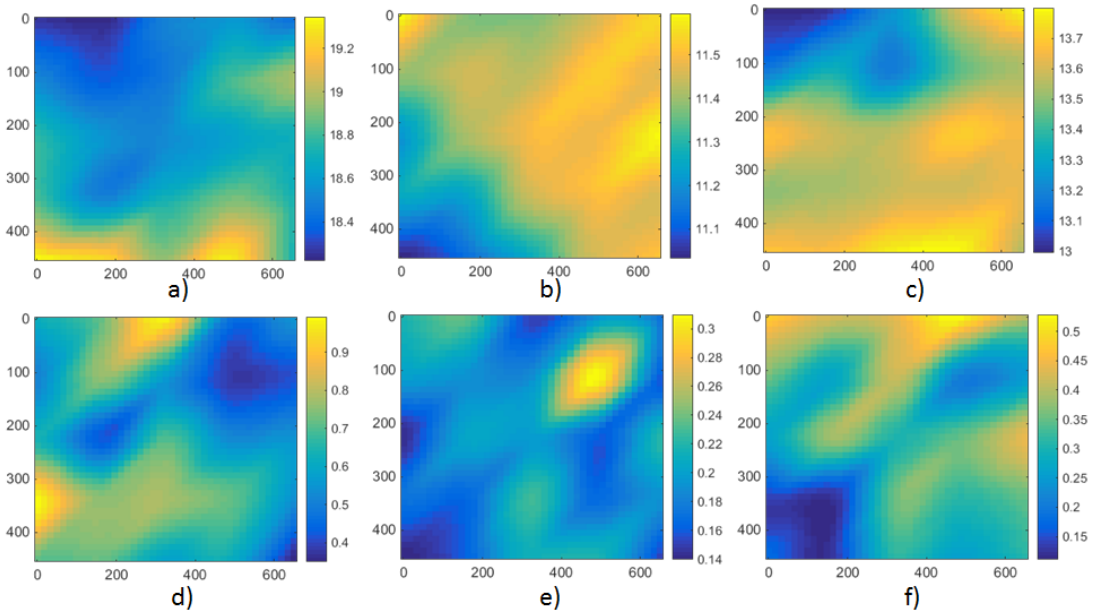

Figure 5.7: Position dependence of out of roundness for 150 µm *dimples, interpolated using bicubic interpolation. a-c) mean out of roundness, d-f) standard deviation of out of roundness. All units in* µm.

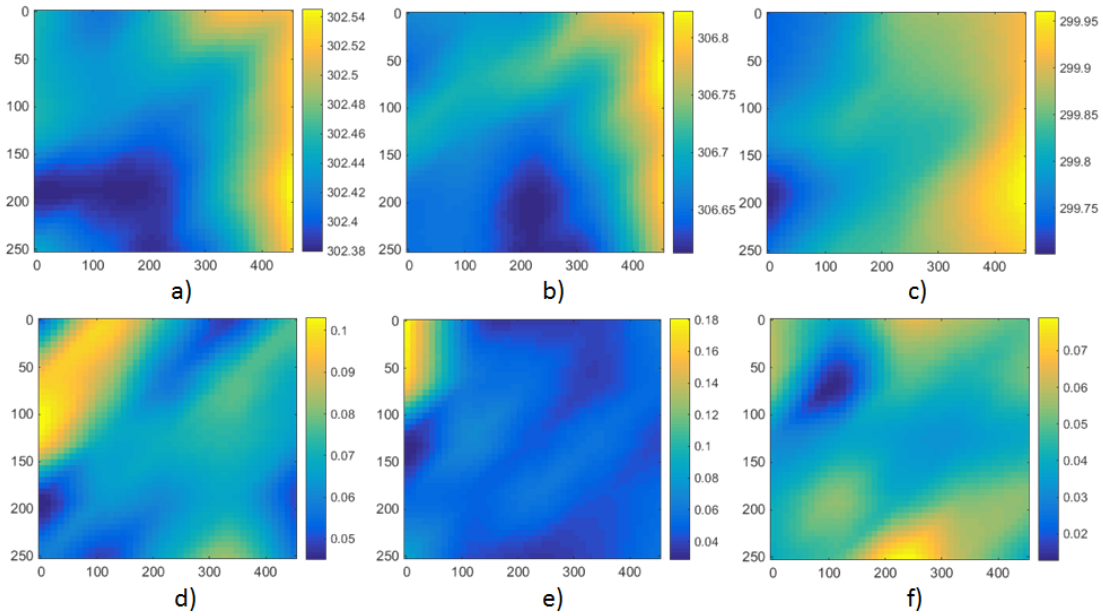

Figure 5.8: Position dependence of diameter for 300 µm *dimples, interpolated using bicubic interpolation. a-c) mean diameter, d-f) standard deviation of diameter. All units in* µm.

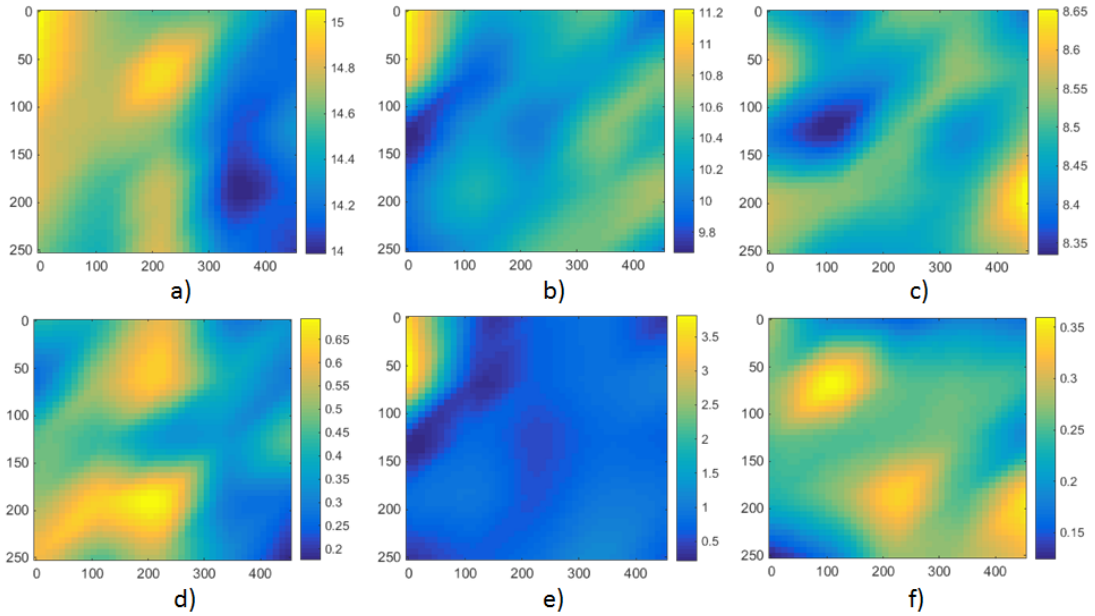

Figure 5.9: Position dependence of out of roundness for 300 µm *dimples, interpolated using bicubic interpolation. a-c) mean out of roundness, d-f) standard deviation of out of roundness. All units in* µm.

Figure 5.10 shows the position dependence of the diameter and figure 5.11 the position dependence of the out of roundness. The height threshold seems to give a fairly consistent position dependence across the three dimples used. This correction could be applied to the data. However, for this analysis no correction was applied due to lack of knowledge about the causes of this distortion.

5.5 Effect of threshold value

Just as the choice of threshold effects the value of geometric properties measured (chapter 4), it is expected that the choice of threshold will effect the uncertainty in these measurements. This issue was investigated by performing the analysis using both gradient and height threshold methods for a range of threshold values. Figure 5.12 shows the resulting uncertainties for the gradient threshold and figure 5.13 for the height thresholds.

Using the gradient threshold, there appears to be an optimal threshold range between $0.2-0.3$, where uncertainty in both diameter and out of roundness are minimized. For the out of roundness, the uncertainty using the gradient threshold varies irregularly between dimples, whereas for the height threshold, the results are far more consistent. Outside these values the uncertainty using gradient thresholds varies significantly between the different dimples. For the height threshold, there is less variability between dimples. However, there is also less of an optimal threshold region. It could be argued that there is lower uncertainty for height thresholds between -0.5 and -1 µm, but the effect is minimal.

These results, highlight the importance of considering multiple features when determining uncertainty, especially if using less stable methods, such as the gradient threshold. There is no consistent relationship between threshold and uncertainty across the dimples, although they all have a similar range of values. Therefore, if only one feature is measured, it is not possible to determine whether the value at a particular threshold is unexpectedly low and may be an underestimate of the uncertainty in measuring similar features.

5.6 Analysis with confocal microscope

A similar repeatability test, using the same measurement strategy as above and analysing the results using the height threshold, was performed using the confocal microscope, although different specific dimples were measured on each instrument. Such a comparison is interesting to see if the instruments have similar capability to measure the dimples

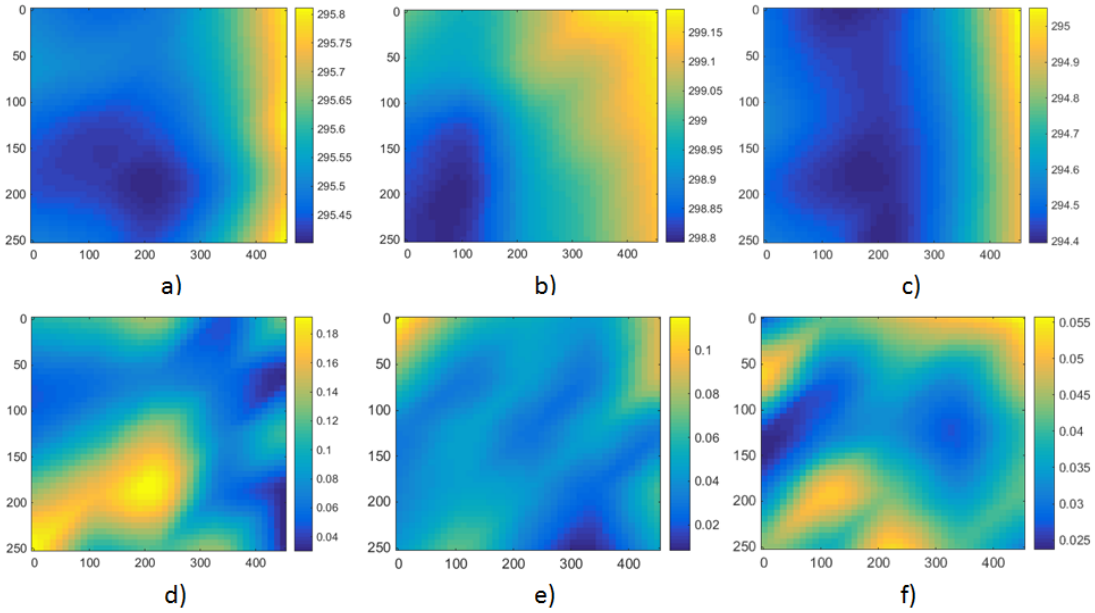

Figure 5.10: Position dependence of diameter for 300 µm *dimples using height threshold, interpolated using bicubic interpolation. a-c) mean diameter, d-f) standard deviation of diameter. All units in* µm.

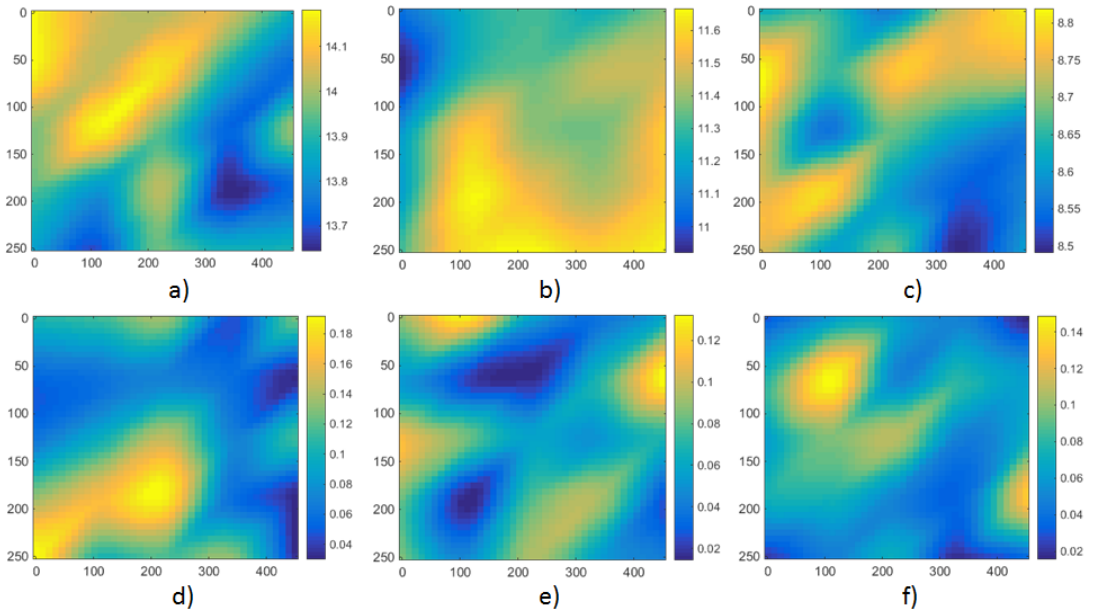

Figure 5.11: Position dependence of out of roundness for 300 µm *dimples using height threshold, interpolated using bicubic interpolation. a-c) mean diameter, d-f) standard deviation of out of roundness. All units in* µm.

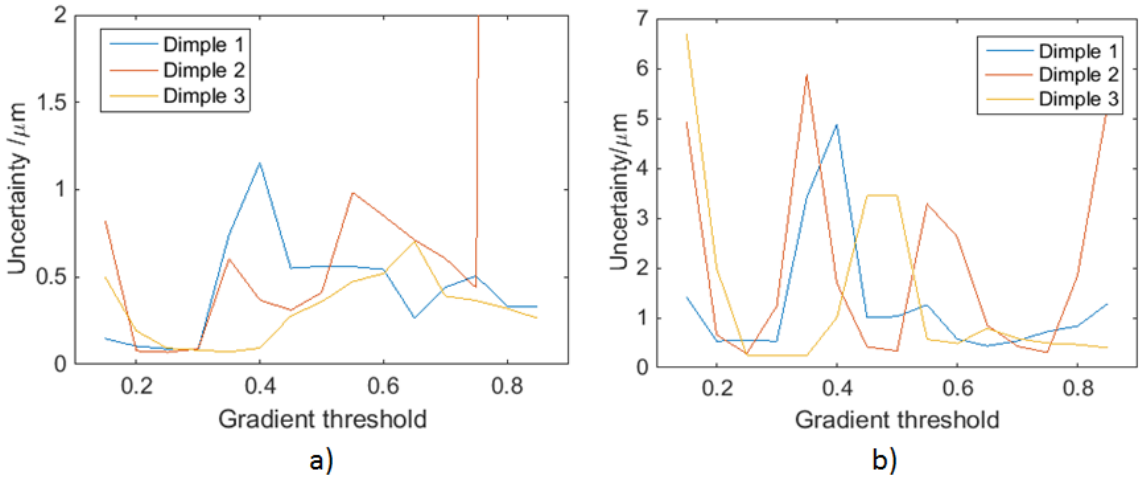

Figure 5.12: Effect of gradient threshold value on standard uncertainty ($k = 1$) for a) diameter and b) out of roundness.

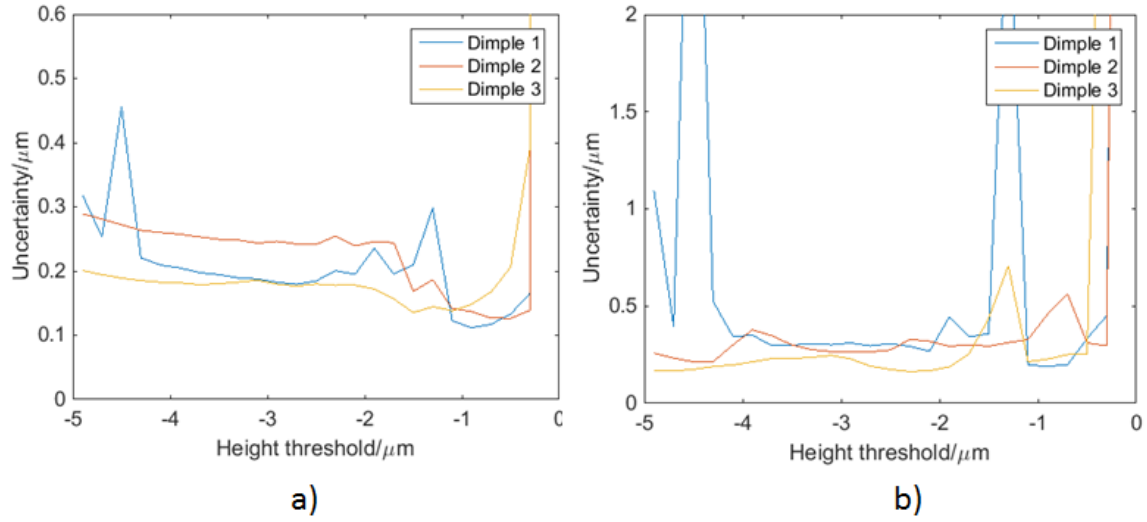

Figure 5.13: Effect of height threshold value on standard uncertainty ($k = 1$) for a) diameter and b) out of roundness.

repeatably, which may be a significant factor in deciding which instrument to use. Additionally, these measurements will be used for comparison to the analysis in chapter 6. The depth is also measured in this case, to assist in that comparison.

Figure 5.14 shows a comparison between the uncertainties for the confocal and focus variation microscopes, using the height threshold. As can be seen, the uncertainty in the diameters are reasonably similar, with the confocal microscope having slightly worse reproducibility. The reason for these differences cannot easily be determined, but one contributing factor may be the additional filtering introduced by the focus variation microscope when calculating the height, which will smooth the topography and reduce the reproducibility. While there may be significant differences due to how the instruments operate, the number of dimples measured may also be an issue. For each instrument, only three dimples out of several thousand on the disk were measured. Without measuring more dimples to allow for estimation of the distribution of the uncertainty for the two instruments it is not possible to determine whether the differences are statistically significant or not.

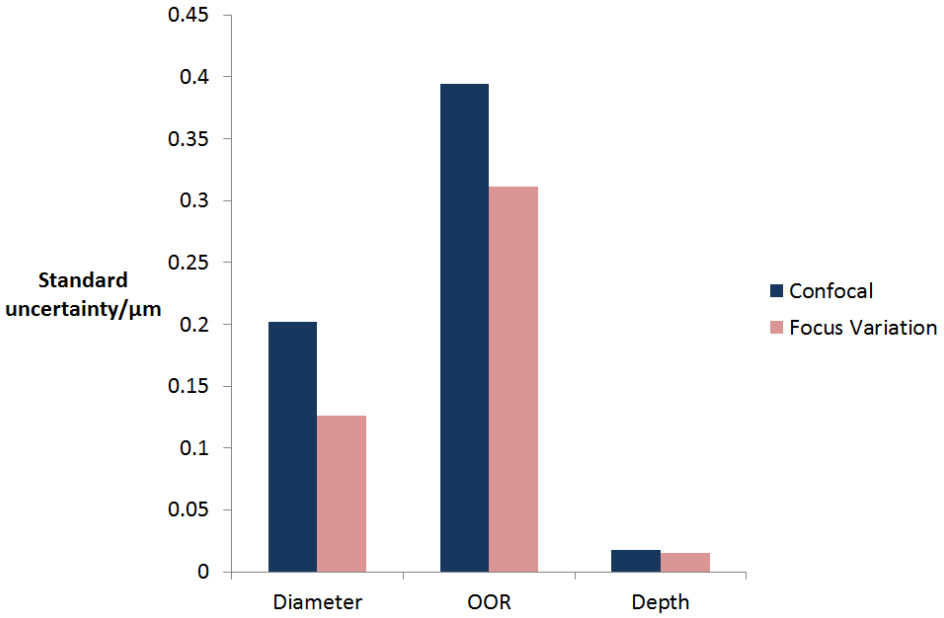

Figure 5.14: Comparison between reproducibility of diameter and out of roundness using confocal and focus variation microscopes

5.7 Discussion

In section 5.2, it was shown that the measurement uncertainty is small compared to the manufacturing reproducibility of the dimples. Therefore, for these samples, the measurement uncertainty will not significantly effect the measurement of the manufacturing reproducibility or the uncertainty of mean values. However, if the surface, manufacturing method or measuring instrument are changed, the measurement uncertainty may have a larger effect. Therefore, it is still important to provide a method to assess uncertainty, in order to show if it will have a significant influence or not.

One of the major limitations of the method presented here is that this method only accounts for uncertainty sources that are transitory in time or dependant on position in the field of view. While this approach accounts for most major uncertainty contributions some, particularly amplification factors due to the traceability of the scales, cannot be considered in this way, as they are expected to be consistent between measurements. If it is believed that these sources are significant, then additional analysis will be required to determine their effect. Similarly, this analysis does not account for any systematic bias in the measurement, which can also be introduced by the same factors that cause uncertainty. In order to consider the measurements as absolute values some way of estimating this bias would be needed, either by estimating the bias and correcting for it, or adding it as an additional uncertainty contribution.

Another issue is the large number of measurements required. In this case, 250 measurements were performed, which for many real applications may be impractical due to the long measurement time. Less measurements could be used but would result in a lower quality estimate of the uncertainty. A simple estimate of the confidence in the uncertainty assumes that measurements are normally distributed and independent. If this is the case, estimates of the variance follow a chi-squared distribution due to Cochran's theorem (Cochran 1934), such that for n measurements with true variance σ^2 and estimated variance s^2

$$(n-1)\frac{\sigma^2}{s^2} \sim \chi^2_{n-1} \tag{5.2}$$

using this relationship for 250 measurements gives an uncertainty in σ such that $0.91s < \sigma < 1.1s$. However, if the number of measurements is reduced to 25, i.e. one measurement at each position with no repeats, this uncertainty increases to $0.74s < \sigma < 1.3s$. In other words, there is an up to 30 % error in the uncertainty estimate if only 25 measurements are performed. In reality, this is an oversimplification, as there is likely to be significant correlation between measurements at the same position and nearby positions on the

grid. However, this approach provides a useful ballpark figure and as a large number of measurements were used it is little interest to provide a better estimate of the confidence in the uncertainty. Although generally, a positive correlation between measurements at the same position, as is observed here, is likely to decrease the overall degrees of freedom and, therefore, result in a larger error in the uncertainty.

Interestingly, there appears to be an optimal range of thresholds for both height and gradient thresholding, which minimises the resulting uncertainty, of -0.5 to -1 µm for the height threshold and 0.2 to 0.3 for the gradient threshold. These lower uncertainty thresholds suggest another approach to deciding a suitable threshold may be to choose the one that minimises the uncertainty. While the uncertainty achieved in using this approach may be lower, it is important to remember that the underlying values will also change and the values with the lowest uncertainty may not be the most physically relevant. Therefore, choosing thresholds and other algorithm parameters solely on the basis of minimising uncertainty is not recommended.

5.8 Conclusion

This chapter has presented an approach to assess uncertainty in structured surfaces via a reproducibility assessment using repeated measurements of the surface. The conclusions are as follows:

1. A reproducibility assessment of the diameter and out of roundness of surfaces with three different nominal diameters was made, using three different dimples on each surface. By repeatedly measuring a grid of 25 positions, uncertainty contributions due to position in the field of view, as well as repeatability, were accounted for. This analysis showed that smaller dimples had a larger uncertainty in diameter, but a smaller uncertainty in out of roundness.

2. Additionally, a larger set of 100 dimples was measured to assess the reproducibility of the manufacturing process. This was much larger than the measurement reproducibility, approximately $3-4$ µm in both diameter and out of roundness. The ratio of variance due to measurement uncertainty over the variance due to manufacturing was calculated and showed that, in this case, the measurement reproducibility will not have a significant impact on the final result, being at most 6.8 % of the variance due to manufacturing and < 1 % when considering the diameter. However, it

is important to remember that for other surfaces the measurement reproducibility may be much larger compared to the manufacturing reproducibility.

3. As well as the value of reproducibility, the effect of position in the field of view was also considered. By plotting the mean and standard deviation of the measurements at each position in the field of view and interpolating between results, maps of the diameter and out of roundness as a function of position were created. While in some cases there appears to be a weak dependence on position, this dependence was not viewed as significant enough to warrant any correction and is just viewed as part of the reproducibility. In other cases, where there is a more significant dependence on position, it may be of interest to investigate further corrections based on the position. However, in such cases, care must be given to not over correct the results, which may introduce greater distortion than was originally present.

4. To show how the segmentation method affects the uncertainty, a comparison was made between the gradient threshold, as was used initially, and a similar height threshold method. The height threshold had a larger uncertainty in the diameter, but a smaller uncertainty in the depth than the gradient threshold. These results reiterate that the choice of method significantly affects the results, as much in terms of uncertainty as in the value itself.

5. Just as the segmentation method affects the uncertainty, it was also shown that the thresholds used affected the uncertainty. Again, the relationship here varied between the height and gradient threshold methods. Interestingly, in both cases there seem to be thresholds which give consistently lower uncertainty than others. This effect is believed to be due to the shape of the topography in these regions and gives the possibility of optimising the segmentation threshold to minimise uncertainty. However, the functional relevance of such a solution would still have to be considered.

Chapter 6

Assessment of uncertainty of geometric properties of structured surfaces using metrological characteristics

Chapter 5 presented an approach to estimate uncertainty in structured surfaces based on repeated measurement of a geometric feature. While this approach is straightforward and relatively easy to calculate, it requires many measurements and cannot easily be generalised between different types of surface, making it time consuming to perform. Additionally, the contribution of the amplification coefficients of the scales cannot easily be assessed, neither can the presence of any bias in the measurements, meaning that reproducibility measurements cannot provide a full assessment of measurement uncertainty.

This chapter investigates an alternative approach to estimate uncertainty using the instrument's metrological characteristics (Giusca et al. 2012a; Giusca et al. 2012b; Giusca and Leach 2013), which estimate the distortions observed in an image in an easily measurable way, and Monte Carlo techniques to propagate the metrological characteristics into the uncertainty. Firstly, the concept of metrological characteristics is introduced and an overview of the measurement strategy is given. Each metrological characteristic used is then considered in more detail, explaining how it is determined and how the results are propagated into an uncertainty. Finally, the results are propagated into a combined uncertainty and various other issues effecting the uncertainty are discussed.

6.1 Set-up and methodology

6.1.1 Sample and measurement set-up

For this chapter, the 300 µm nominal diameter, 10 µm nominal depth, 20 % texture density disk is used for the analysis; as the aim of this chapter is to demonstrate the analysis method used, it is not necessary to assess the full range of diameter, depth and texture density available. However, the approach used here would be just as applicable to any of the other disks, or other structured surfaces.

Measurements for this chapter were made using the confocal microscope, described in section 3.2. In theory, this approach could be applied to any other instrument, provided that its metrological characteristics can be determined. However, focus variation microscopy was not used here, because the relatively high roughness required to perform focus variation measurements makes determining the instrument's metrological characteristics difficult. Many of the standard artefacts used to determine the metrological characteristics are too smooth to use with focus variation instruments. It is possible to estimate the characteristics of such instruments using specially roughened samples and a large number of repeat measurements. However, such measurements are challenging and research into the optimal approach is currently ongoing (Giusca et al. 2014).

For this chapter, only two geometric properties of the feature are assessed, diameter and depth. Out of roundness which was considered in chapter 5 is not considered here. In theory, any other properties could be considered using a similar approach to the one presented in this chapter, provided that care is taken to ensure the full variation of these parameters is simulated. The difficulty lies in ensuring that suitable variation in parameters is measured, particularly for parameters which are dominated by local effects, such as out of roundness, where additional effects such as the rotation of a feature in the field of view may have a significant effect. Therefore, it was decided to demonstrate this method only using the depth and diameter which are dominated by the average effects over the image reducing the amount of variation that must be considered. Extending this method to other parameters is left as future work.

6.1.2 Characterisation approach

For this chapter, the same height thresholding algorithm as used in section 5.4 was used to segment the surfaces. However, the approach to determine uncertainty is not reliant on the segmentation algorithm and the same approach can be used for any algorithm

that produces a binary segmentation of feature and background regions. Just as in chapter 5, the uncertainty will vary with choice of segmentation algorithm, as how changes in the topography, caused by the metrological characteristics, effect the segmentation is controlled by the algorithm used.

6.1.3 Metrological characteristics

Uncertainty in a measurement is determined by many influence factors of an instrument such as, scanning non-linearities, optical distortions and quality of light source, which can effect the measurement results (Leach et al. 2015). However, in practise it can be difficult to measure these influence factors and how they affect a measurement, especially for commercial instruments where some factors are not known or are not easily measured. Additionally, the relevant influence factors and their effects will vary depending on the instrument used.

One method to avoid these problems is to only consider the output of the measurement, treating the internal workings of the instrument almost like a black box. The errors in the instrument can then be characterised by the so called metrological characteristics, which described the geometric distortions observed in a measurement (Giusca et al. 2012a; Giusca et al. 2012b; Giusca and Leach 2013; Leach et al. 2015). The relevant metrological characteristics for areal surface topography instruments are currently being standardised in ISO 25178-600 (ISO 2015). This document is currently in the draft stage and so is subject to change, but it describes a set of seven metrological characteristics which describe the main errors in surface topography instruments. These are:

- Amplification coefficient (vertical and lateral)
- Linearity errors (vertical and lateral)
- Flatness deviation
- Measurement noise
- Spatial resolution of measurements
- Perpendicularity of axis
- Topography fidelity

There are various standard methods that can be used to determine these parameters. The methods used in this analysis are described below. However, other methods would be equally valid, see Giusca et al. 2012a; Giusca et al. 2012b; Giusca and Leach 2013; ISO 2015; Giusca and Leach 2012 for more information on metrological characteristics and how to determine them.

In this chapter, the measurement noise, flatness deviation, vertical and lateral amplification coefficients and linearity errors are measured and used to estimate the uncertainty in the measurement of a dimple. For various reasons the other metrological characteristics are not measured for the reasons discussed below.

The perpendicularity of the vertical axis is rarely assessed for commercial surface topography instruments, as it is assumed to be small and have a negligible effect (Giusca et al. 2012b). However, in the xy plane the perpendicularity of the axis is indirectly assessed and is included in the measurement of linearity errors, so further assessment is unnecessary.

For the resolution, the vertical resolution is generally limited by the measurement noise and can be ignored. However, lateral resolution can have a significant effect on measurements. Lateral resolution can be assessed using a star artefact or similar (Giusca and Leach 2013), which gives a lateral period limit, where the response of the instrument drops below 50 %. However, the lateral resolution assumes that the instrument response is linear, which is often not the case, particularly when measuring steep slopes. To deal with this issue the topography fidelity was introduced, which aims to describe the errors measured in a traceable surface topography, including non-linear effects. However, as the topography fidelity is still a new characteristic, the best way to assess it is not clear; especially as it includes non-linear effects, which will be dependant on local topography. For these measurements, an initial comparison was performed with stylus measurements, as shown in figure 6.1. This comparison indicates that effect of the lateral resolution is suppressed by the relatively large filter, of approximately 10 µm, compared to the lateral resolution of approximately 2 µm. Therefore, any distortions introduced by the lateral resolution, will be insignificant compared to the effect of the filter and can be ignored. Similarly, it is believed that the distortions in the topography fidelity will be also suppressed by the filter used. While it is not generally true that topography fidelity can be suppressed by filtering, it is believed to be the case for these measurements.

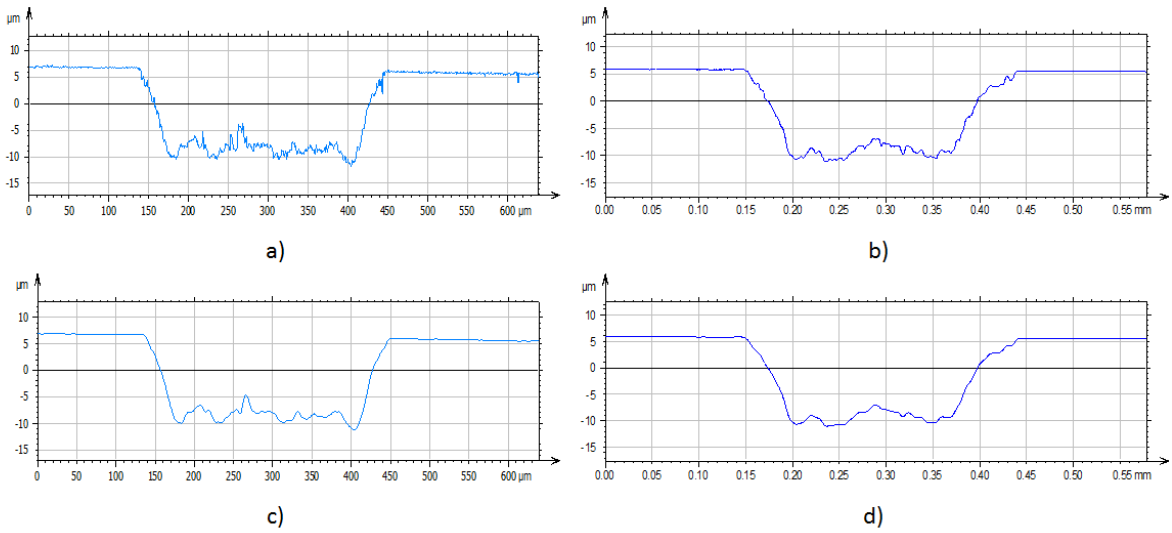

Figure 6.1: Comparison between stylus profiles and profile from confocal measurement. After filtering the results are similar with each method. Note different features are measured in each case. a) Profile extracted from confocal measurement of surface. b) Stylus measurement of a feature. c) Confocal measurement after filtering with a 8 µm *Gaussian filter. d) Stylus measurement after filtering with a* 8 µm *Gaussian filter.*

6.2 Assessment of uncertainty

To determine uncertainty in the geometrical characterises of a feature, here diameter and depth, the effect of the metrological characteristics must be propagated through the analysis. The steps required to propagate these uncertainties vary depending on the metrological characteristic being assessed and the process for each is discussed below.

To perform the analysis, the same three dimples were used as in chapter 5, allowing for easier comparison between the two methods of assessing uncertainty. For this analysis, only a single measurement of each dimple is needed and is combined with the assessment of the metrological characteristics as described below.

6.2.1 Measurement noise

The measurement noise is defined as the expected contribution from random noise between repeated measurements under the same conditions and can be estimated using repeated measurements of a surface (Giusca et al. 2012a). In this case, a subtraction approach was used where the difference between pairs of measurements is used to determine

noise. The noise, ϵ, at a particular pixel from a pair of measurements is given by

$$\epsilon = \frac{z_1 - z_2}{\sqrt{2}} \tag{6.1}$$

where z_i is the ith height measured at that pixel. This can be used to determine a distribution of the noise over the entire surface. As optical surface topography instruments are sensitive to factors such as local slope and material effects, the measured noise is dependent on the surface topography. Therefore, measurements should be taken on a surface that closely matches the surface of interest. In this chapter, 12 repeat measurements of a randomly chosen dimple were made and noise was assessed on subsequent pairs of measurements, i.e. $z_1 - z_2$, $z_2 - z_3$, etc. The noise distribution was then calculated based on the combined result of all these pairs, in order to provide the best noise estimate.

Normally, noise is estimated by the root mean square noise of the surface (Sq_{noise}) and for simplicity an uncorrelated Gaussian distribution over the surface is often assumed (Giusca et al. 2012a; Leach et al. 2015). However, for structured surfaces, the noise values often differ significantly between different regions of the surface; in this case background and feature regions, as shown in figure 6.2. In such cases, a more in depth analysis, differentiating between the different regions is required. This can be achieved by segmenting the surface using the same approach as when analysing the dimple's properties and then calculating noise in each region using the pixels in those regions.

Additionally, the noise, shown in figure 6.2, has a non-Gaussian distribution, with many large spikes indicating a more heavy tailed distribution. Therefore, in this case, it was decided to use a bootstrap of the noise distribution directly, rather than using the Sq_{noise} values as would normally be done. This means that the noise map was segmented into feature and background regions and for each pixel in the measurement a random value from the appropriate region in noise map was chosen (with replacement) and added to the measured value to simulate the noise. Such an approach may not be necessary for other instruments depending on their noise distributions. However, the Sq_{noise} still provides a convenient indication of the magnitude of the noise, which is more difficult to understand directly from the distributions. For the dimple in figure 6.2, the average Sq_{noise} was measured as 170 nm for the background surface and 460 nm in the dimple.

To estimate the effect of measurement noise, a Monte Carlo simulation was then performed, where to each pixel random noise was added based on the appropriate distribution, either inside or outside the feature, before analysing the surface to determine the diameter and depth of the feature. This process was repeated 5000 times, using different random noise values to build up distributions, shown in figure 6.3 for one of the dimples

used. In this case, the 95 % confidence intervals are $294.33 < \phi < 994.62$ µm for diameter and $7.627 < d < 7.635$ µm for depth. The full results are summarised in section 6.4.

Noise may also introduce a bias into the results depending on how it interacts with the segmentation algorithm. This can be determined by calculating diameter and depth results for the surface with no noise added and comparing them to the mean of the distributions in figure 6.3. This gives a bias due to noise of 80 nm in the diameter and 12 nm in depth, where a positive value the measured diameter is larger than the true value.

6.2.2 Flatness deviation

Systematic shifts in the height across the field of view can also effect measurements, due to optical distortions or other effects. These effects are termed flatness deviation and commonly cause a flat surface to appear bowed. To estimate the flatness deviation, a flat artefact is measured. It is important to remove any residual topography in the artefact and minimise the effect of measurement noise, which could mask the flatness deviation. Therefore, the sample should be measured multiple times, with a small shift in position between measurements and the results averaged (Giusca et al. 2012a). For this chapter,

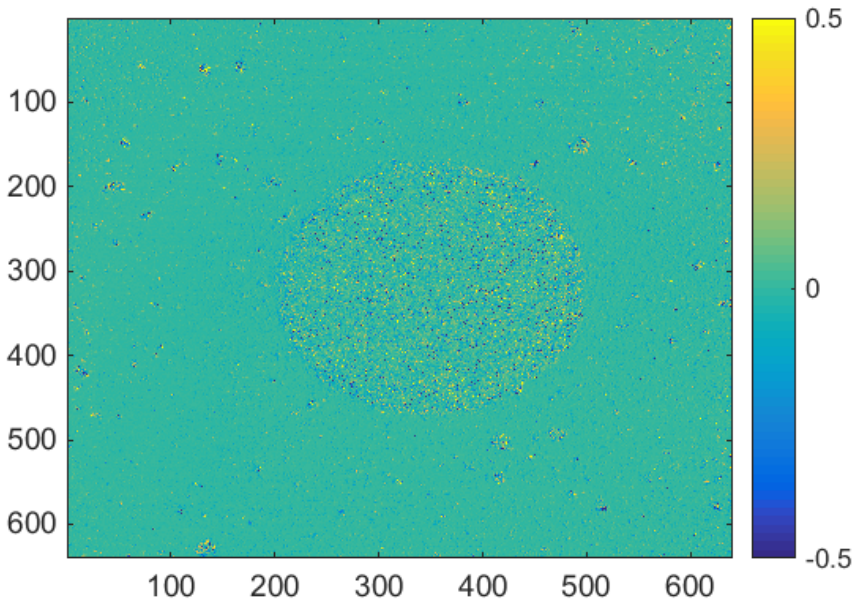

Figure 6.2: Noise map produced by the difference between two measurements. All axis in µm. *z axis has been truncated to increase clarity, full range is approximately* ±8 µm.

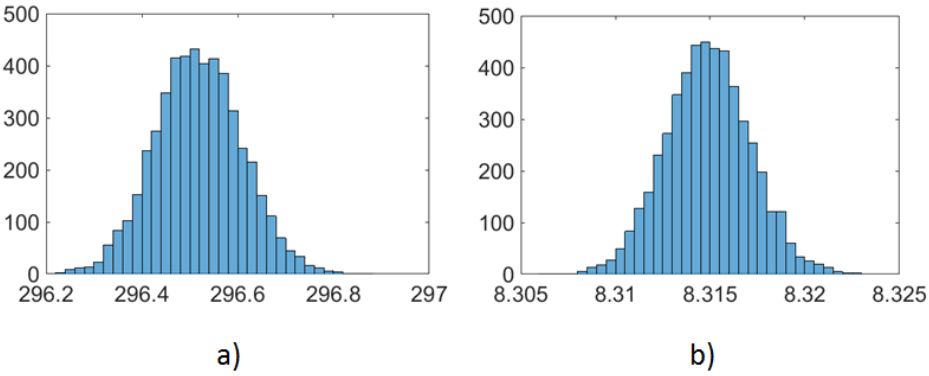

Figure 6.3: Histogram of Monte Carlo simulation results for a) diameter and b) depth. x axis in µm*. y axis is in counts.*

a 5×5 grid of measurements was taken with 100 µm shift between measurements and a flatness map was calculated as the mean of these measurements after levelling and outlier removal.

Outlier removal is necessary as outliers in the data, due to measurement noise or deformations in the surface, can significantly affect the flatness measurement, even after averaging. Therefore, some method to remove these outliers is needed. Several possible methods to perform this outlier removal exist. In this chapter, form was first removed by fitting a 5th order polynomial to the surface. A 5th order polynomial was found to give good results in this case, although other choices of polynomial order are also valid (Giusca et al. 2012a). Then, pixels with heights more than three standard deviations from the mean of the residual surface were marked as outliers and not added to the averaged image. It should be noted that the form was only removed for the purpose of determining outliers and the original topography (excluding outliers) was used when calculating the average map.

Normally, the flatness deviation is reported as a Sz value, i.e. peak to valley deviation (Giusca et al. 2012b; Leach et al. 2015). However, for the analysis here, the Sz value is not sufficient, as it gives no indication of the shape of the flatness deviation, and the map of the flatness deviation, shown in figure 6.4, is used instead. As figure 6.4 shows, for the confocal microscope there is a significant flatness deviation in the instrument of approximately 200 nm. By performing adjustment routines on the instrument it may be possible to improve this value. However, such a process is beyond the scope of this chapter and is not considered here.

The effect of the flatness on the final geometrical properties will depend on the position

of the feature in the field of view. One way to assess the contribution of the flatness deviation is to subtract the flatness map from the measurement data, i.e. correct for the effect of flatness, then add the flatness map back to the measurement after a shift in x and y. This process simulates performing the same measurement at different positions in the field of view. Figure 6.5 illustrates the method used.

To determine the uncertainty contribution of the flatness, the above process is applied via a Monte Carlo simulation using 5000 repeats with a range of shifts. The position of the feature in the field of view and, therefore, the shift was modelled as a uniform distribution with as large as possible range to avoid the feature clipping the edge of the image. In reality, at least for manual positioning, central positions are probably more likely. However, bigger distortions are expected at larger shifts. Therefore, a uniform distribution provides a conservative uncertainty estimate.

Figure 6.6 shows histograms for one of the measured dimples from the Monte Carlo simulation. These results give the 95 % confidence intervals as $294.42 < \phi < 294.83$ µm for the diameter and $7.619 < d < 7.698$ µm for the depth.

The flatness may also introduce a bias into the results. As before, this can be determined by calculating diameter and depth results for the surface with the flat corrected and comparing them to the mean of the distributions in figure 6.6, with the distributions shown giving a bias due to flatness of −416 nm in the diameter and −76 nm in depth.

6.2.3 Z scales

The amplification coefficient and linearity errors in z scale can be assessed together, using measurements of a series of calibrated step heights (Giusca et al. 2012b). For this chapter, six different step heights with nominal steps of 350 nm, 500 nm, 1.2 µm, 2.4 µm, 3.0 µm and 17 µm were used. Several factors should be considered to get useful measurements from these steps.

Firstly, there may be a systematic change in the amplification as a surface is moved up or down in the measuring volume of the instrument. To account for this effect, three sets of step measurement are taken, one at a normal measurement position (0 mm) and additionally where the position of the surface in the measurement volume is changed by ±1 mm. For the confocal instrument used this step can be achieved by moving the focus position position up or down by 1 mm and adjusting the manual (course) focus so that the surface is in focus, but for other instruments it may be necessary to physically raise or lower the surface. Additionally, it is important to assess the reproducibility of

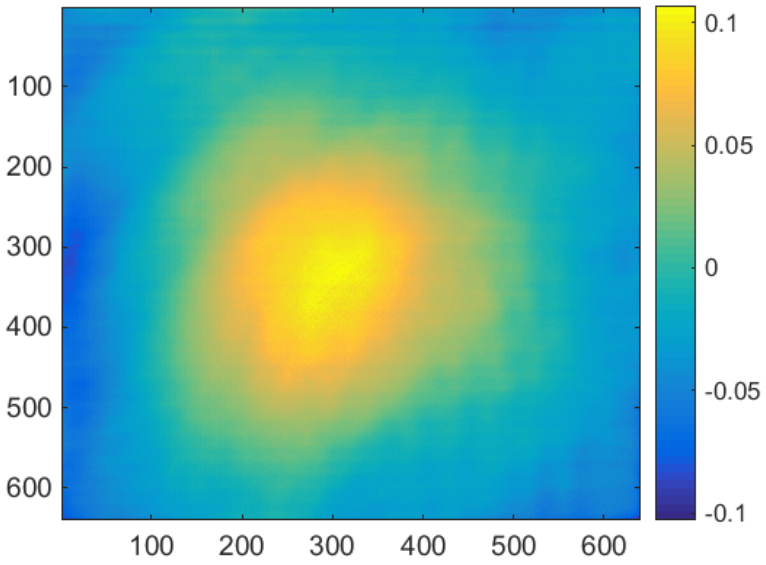

Figure 6.4: Flatness map calculated from 25 repeats. All units in µm.

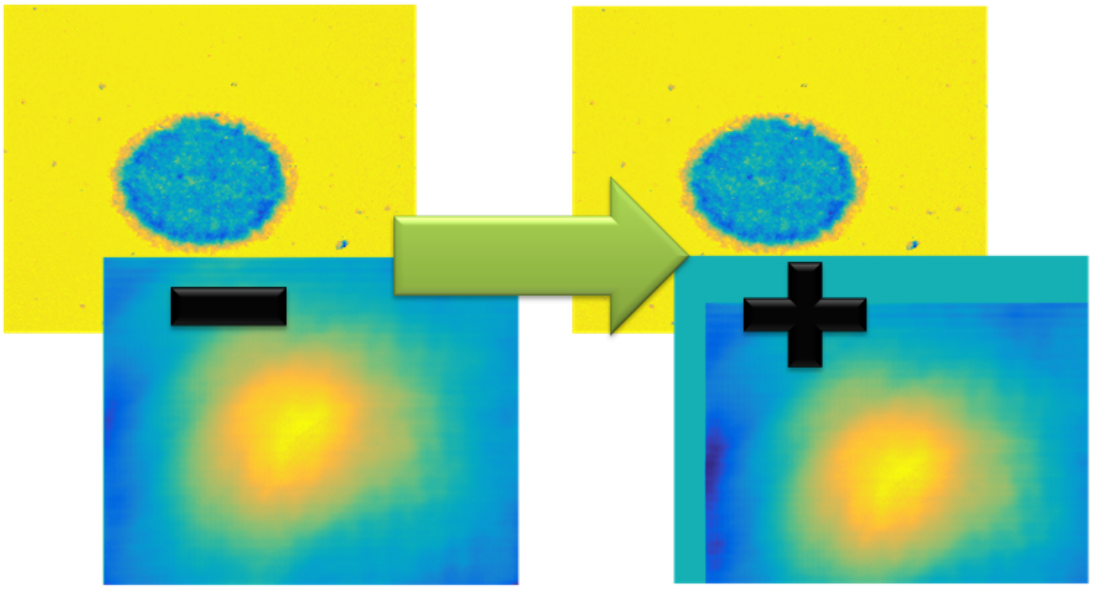

Figure 6.5: Method to determine effect of flatness. First flatness map is subtracted from the measured data. Then flatness map is shifted and added back to height data. Geometric parameters are calculated on resultant height map.

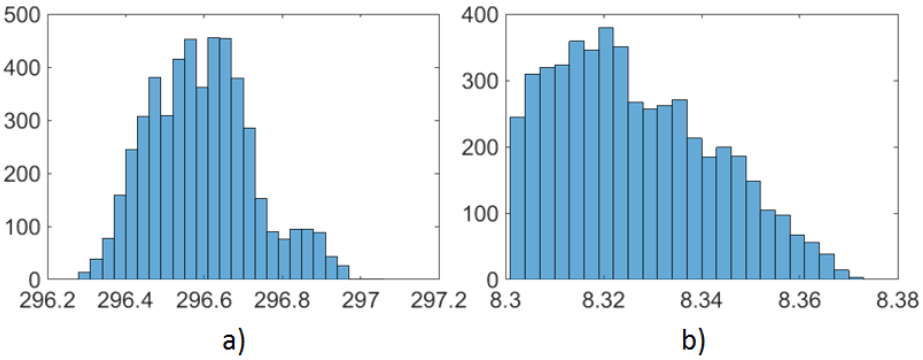

Figure 6.6: Histogram of Monte Carlo simulation results for a) diameter and b) depth. x axis in µm*. y axis is in counts.*

the step height measurement, in order to determine its uncertainty. Therefore, five repeat measurements of each step, at each position where made, for a total of 15 measurements of each step.

Once completed, the measurements showed no significant dependence on position in field of view and changes between sets of measurements of the same step are viewed to be due to the non-linearities in the system. Therefore, the results could be combined to give a single mean and uncertainty, determined from the standard deviation of the measurements, of each step. These results, along with the calibrated heights are shown in table 6.1. A linear fit through the origin, shown in figure 6.7, can then be used to estimate the amplification coefficient, as 0.9989 ± 0.0032 ($k = 2$).

Table 6.1: Nominal, calibrated and measured heights and expanded uncertainties ($k = 2$). Uncertainties in measured values from repeatability of the measurements.

Nominal height/nm	Calibrated height/nm	Measured height/nm
350	343.2 ± 2.3	331.2 ± 3.8
500	500.4 ± 4.1	460.8 ± 38.6
1200	1267.4 ± 4.2	1205.6 ± 16.0
2100	2055.0 ± 4.1	2040.0 ± 35.4
3000	3030.5 ± 4.3	3012.9 ± 8.8
17000	17018.2 ± 26.9	17009.5 ± 25.6

As well as the amplification factor, an estimate of the non-linearities must be made. A simple estimate can be given by the residuals of the linear fit, shown in figure 6.8. Due to

small number of sample heights, the effect of linearity error is modelled as a rectangular distribution with bounds at ±|max error|. While this approach likely overestimates the linearity error, without further information on how the residual varies with height it is not possible to make a better estimate. One possible way to achieve more information on the non-linearity may be by using a tilted flat (Kiyono et al. 2000). However, this approach is non-trivial, requiring a separate interferometer to measure change in height, and was not attempted here. Based on the results in figure 6.8, the linearity error of the system is estimate to be ±72 nm.

This non-linearity is surprisingly large. Further analysis shows that some of the 500 nm, 1200 nm and 2100nm step measurements have a large difference from the calibrated value, as can be seen in figure 6.8. These three steps were made of nickel whereas the other steps, where this effect did not occur, were glass. These nickel steps appear to result in an error where the background surface is shifted, as shown in figure 6.9. Whether this error is caused by the material, the square shape of the steps or some other effect is currently unknown. However, the brightness setting of the confocal instrument is believed to be a factor and reducing the illumination when performing measurements can significantly reduce the distortion to almost zero. However, such conditions do not represent the real measurement conditions when measuring a feature and it is not known if this effect occurs when measuring the dimples. Therefore, only the original results were used, even though this may be an overestimate of the real effects.

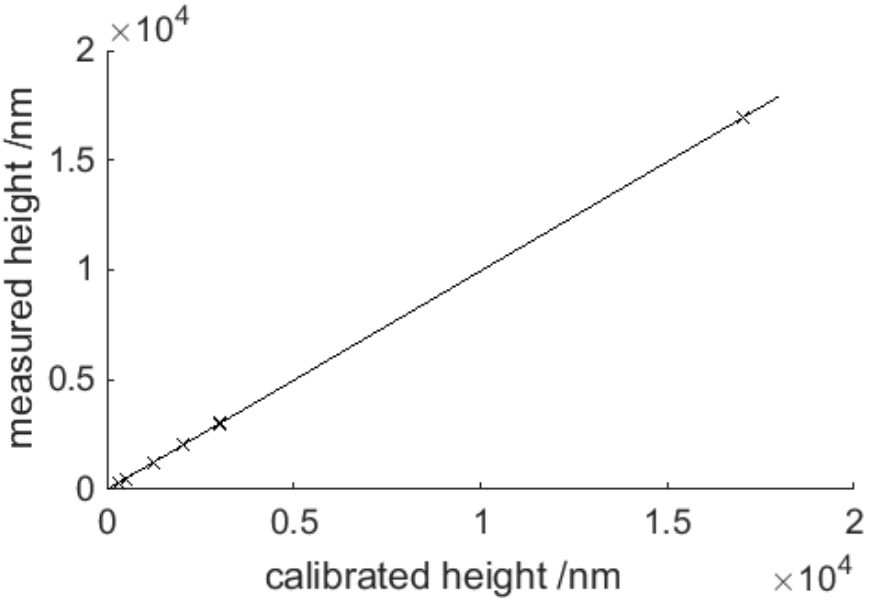

Figure 6.7: Plot of calibrated and measured step height values.

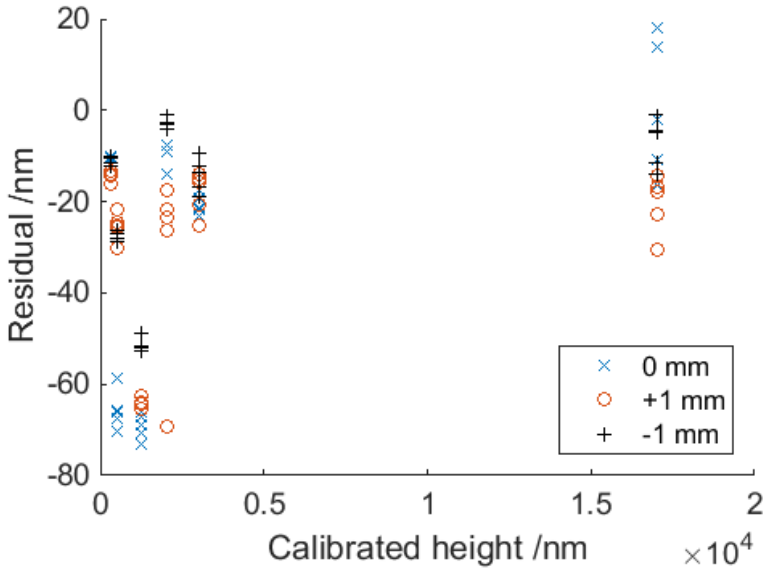

Figure 6.8: Residuals after removing mean amplification coefficient from measured data.

The effect of the z scales on dimple depth and diameter was again assessed using a Monte Carlo approach. The non-linearity was evaluated as a simple modification of the amplification factor over the nominal depth of the feature (10 µm). While this is a gross over-simplification of the real effect of the non-linearity, without more information

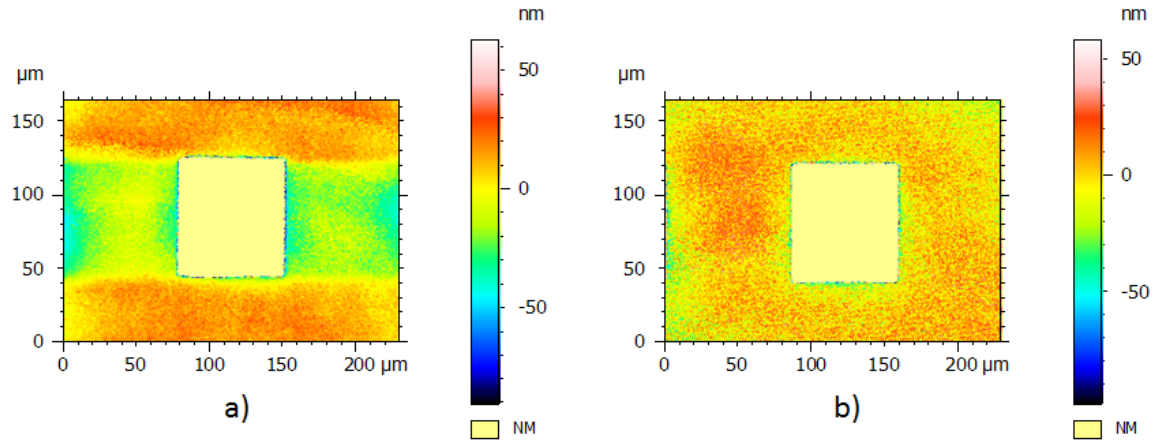

Figure 6.9: Error caused by measuring nickel steps. Measurements have been thresholded to highlight the errors. a) Distortion in the background surface can be seen in the measurement of the step. b) When measured using a much lower brightness setting the distortion is eliminated.

on how the non-linearity changes with height, a more realistic estimate is not possible. Random estimates of the amplification coefficient and non-linearity are then made, giving a modified amplification coeffcient α_z modelled by

$$\alpha_z \sim N(\alpha_{z(lin)}, u_\alpha) + \frac{U(-|\text{max error}|, +|\text{max error}|)}{10000} \tag{6.2}$$

where u_α is the standard uncertainty in the amplification factor and $\alpha_{z(lin)}$ is the estimated amplification factor before modification. The height value of the measurement was then modified by this amplification factor and values of diameter and depth calculated. By repeating this process 5000 times, distributions of diameters and depths were built up, as shown in figure 6.10 for one of the dimples used. The corresponding 95 % confidence intervals of this distribution are $294.38 < \phi < 294.42$ µm for the diameter and $7.573 < d < 7.648$ µm for the depth.

Just as for the other parameters, the z scale may also introduce a bias in the results. By comparing the diameter and depth calculated with amplification factor 1 with the means of the distributions in figure 6.10, the biases are determined to be 15 nm in diameter and −8 nm in depth.

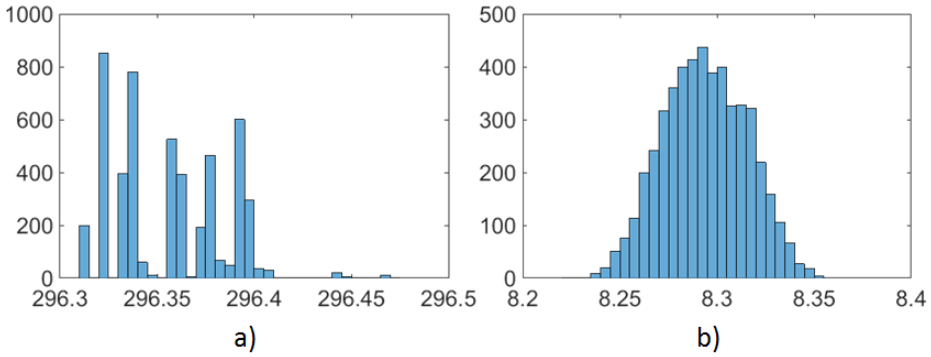

Figure 6.10: Histogram of Monte Carlo simulation results for a) diameter and b) depth. x axis in µm*. y axis is in counts.*

6.2.4 Lateral scales

Errors in the lateral (xy) scales can be assessed separately from the vertical scales using a cross grating artefact (Giusca et al. 2012b), as shown in figure 6.11. The centre of mass of the cross-grating pits, extracted by height thresholding can then be used to determine non-linearities in the lateral scales.

One approach to assess the lateral distortions, which was applied in this chapter, is self-calibration (Henning et al. 2013; Downs et al. 1999). By measuring the artefact three times including a shift and rotation, any distortion in the scales can be separated from errors in the artefact. In this way, a distortion map over the field of view can be created.

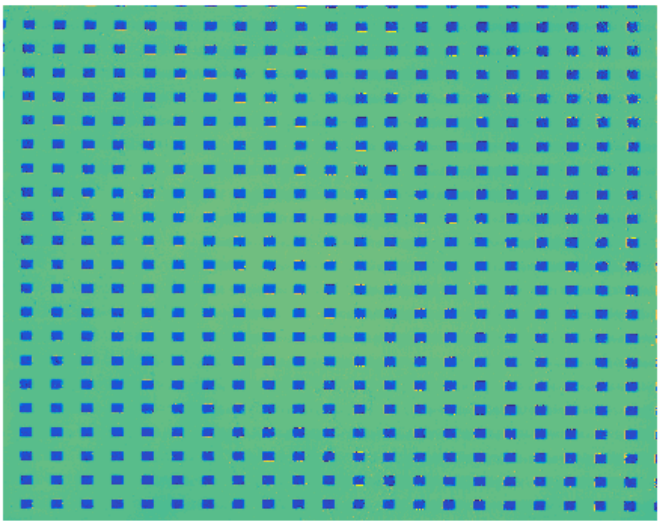

Figure 6.11: Height map of cross grating. Note errors on grid edges these are not believed to be in the artefact but due to instrument effects on the edges.

For these measurements, it was found that there was significant non-repeatability in determining the pit centre of mass, due to various factors including pixel size and edge effects at the pit boundary. Therefore, five measurements were taken at each position, repositioning the grid between each measurement. The centres were then extracted and aligned via a rigid translation to account for the repeatability of the stage. The mean at each grid position was then taken as the centre of mass at that position.

The result of these measurements was three sets of grid coordinates, one set rotated by 90° compared to the first and another set shifted by one grid row in the x direction compared to the rotated measurements. A self-calibration approach (Henning et al. 2013; Downs et al. 1999) was then used to determine the distortion of the grid points in x and y. Figure 6.12 shows the distorted and nominal grid points showing 10 times the real error. These errors were interpolated using a bi-cubic interpolation scheme to produce maps of the x and y errors, shown in figure 6.13. The self-calibration algorithm also gives an uncertainty in the error at each point. The value changes over the field of view but an

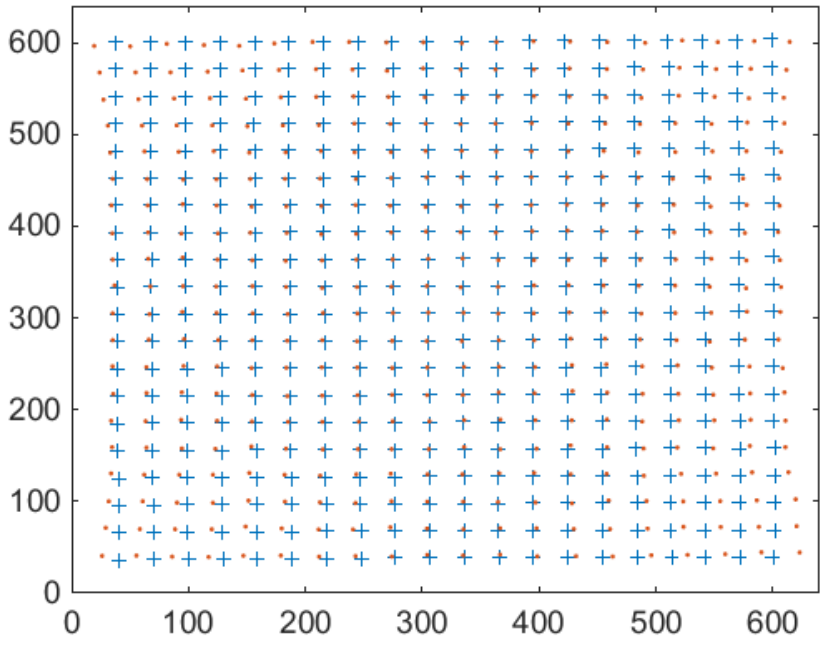

Figure 6.12: Nominal grid positions (blue crosses) and distorted positions (orange dots) from self calibration. Distortion is magnified magnified by a factor of 10 to allow easy visualisation. Box shows entire field of view with axis in µm.

average value of 0.3 pixels or 188 nm in each direction will be used in further analysis.

From the error map, a similar approach as for the flatness deviation was used. First, the edge of the measured feature was extracted. Distortion in the edge was then corrected by subtracting the x and y distortions at the position of the measured edge. The distortion was then added back after shifting the distortion map. The uncertainty due to the distortion was determined via a Monte Carlo simulation using 5000 random shifts, with the diameter and depth of the feature calculated at each point.

Figure 6.14 shows histograms of the result of the Monte Carlo simulation on one of the dimples used. These results give the 95 % confidence intervals of $294.04 < \phi < 294.29$ µm for the diameter and $7.5305 < d < 7.5316$ µm for the depth.

Uncertainty in the lateral scale distortion was accounted for by repeating the analysis at a number of measurement positions varying the position of each boundary pixel, using a 0.3 pixel standard deviation Gaussian. In reality, nearby points will be somewhat correlated, but considering variations between pixels as uncorrelated provides a good initial estimate and, if this uncertainty is believed to be a significant contribution to the

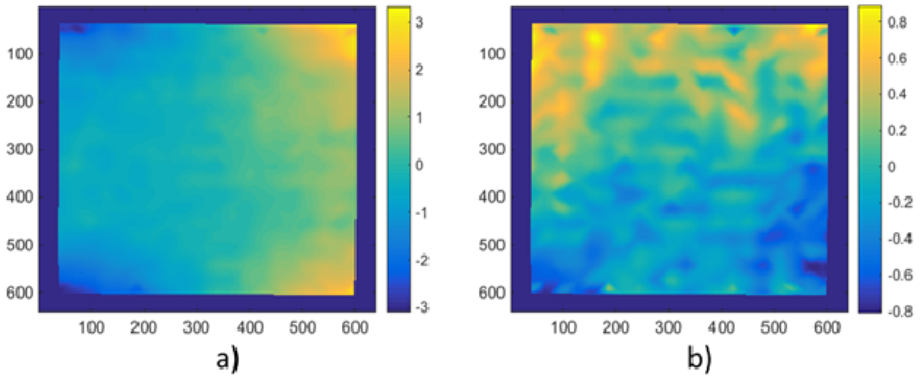

Figure 6.13: Distortions in x (a) and y (b), interpolated using a bi-cubic interpolation. Images show entire field of view with all axis in µm.

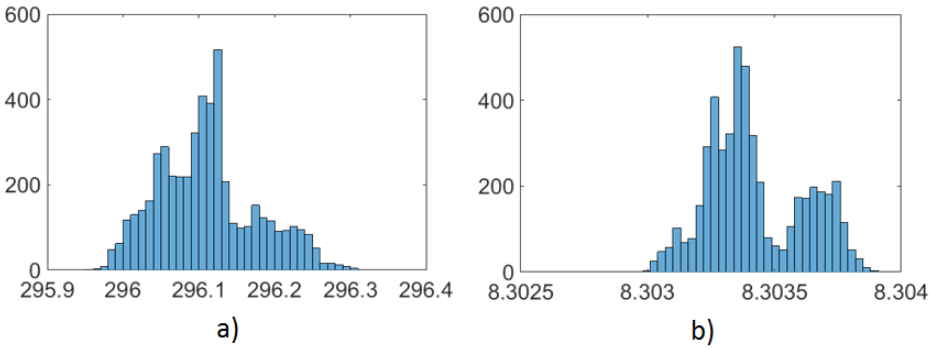

Figure 6.14: Histogram of Monte Carlo simulation results for a) diameter and b) depth. x axis in µm*. y axis is in counts.*

overall measurement uncertainty, further analysis to account for any correlation can be performed. However, in this case, the uncertainty in the lateral scale distortions results in a 56 nm standard uncertainty in diameter and < 1 nm in depth, which is much smaller than some other contributions and only has a minor effect on the combined uncertainty.

The lateral distortion may also introduce a bias in the results. By comparing the undistorted diameter and depth with the mean values from figure 6.14 the biases are -51 nm in diameter and < 1 nm in depth.

The above analysis only considers non-linearities in the lateral scales; the amplification should also be assessed. Ideally, a calibrated grid would be used and a direct comparison between the calibrated positions and measured positions, after the non-linearities are corrected, would allow for the amplification factor to be determined. With a calibrated grid, the uncertainty will be limited by the uncertainty in the calibrated grid,

with uncertainties of 100 nm being achievable over a field of view (Giusca et al. 2012b). Unfortunately, a suitably calibrated grid was not available for these measurements and the grid, in figure 6.11, is uncalibrated. While this does not effect the non-linearity estimate, as it uses a self-calibration approach, it is not possible to estimate the amplification without a known distance. Another possible approach would be to use another calibrated artefact with a known distance, if available. A lower quality estimate is also often given by the manufacturer's calibration of the instrument, which is used here. The calibration certificate of the instrument gives the maximum lateral error as 1.03 %. However, this value also includes non-linear effects already accounted for by the grid measurements. Indeed, a 1 % error over half the field of view is approximately 3 µm, the same as the maximum observed lateral distortions, suggesting that the actual amplification in the lateral scales is relatively small.

6.3 Filter effect

As shown in chapter 4, the choice of filter size can have a significant effect on the measured geometric properties of a feature. As the filter size is based on the nominal pixel spacing, any lateral scale effects will have an influence on the real filter used. Therefore, the uncertainty in the lateral scales will introduce uncertainty in the filter size used and will add additional uncertainty to the geometric properties. For this example, only the 1 % uncertainty in the amplification factor contributes to the uncertainty using a linear model, as described in section 2.8.5, where the diameter and depth of the dimple are redetermined after changing the filter size by ±1 %. The uncertainty is then modelled as a uniform distribution between the maximum and minimum values. For the example dimple used previously, this model gives a change in diameter of 44 nm, corresponding to a standard uncertainty 13 nm. For the depth, the effect is very small with a change of < 1 nm.

This sort of linear model is commonly used to model uncertainty (see section 2.8.5 or JCGM 2008a; Haitjema 2013). However, it relies on the assumption that the effect being modelled behaves approximately linearly over the range of variation, which the results in chapter 4, show is not the case. A more robust approach would be to perform a Monte Carlo simulation using random changes in the amplification factor to build up the distributions. However, as the effect of filtering in this case is relatively small compared to other effects, a more robust simulation was not deemed to be necessary and a simple estimate was sufficient.

6.4 Combined uncertainty

To estimate the combined uncertainty in a measurement of the diameter or depth of a dimple, the uncertainty due to all the previously mentioned sources must be combined. Two approaches to combine the uncertainties are considered. Either using a Monte Carlo approach, drawing random values from each distribution and combining them to build up a combined distribution, or relying on the central limit theorem and combining the standard deviation of each uncertainty source in quadrature as per the law of propagation of uncertainty in the GUM (JCGM 2008a). Due to the non-Gaussian nature of some of the uncertainty terms, the Monte Carlo approach was initially used. Figure 6.15 shows the resulting distributions for one of the dimples used. This gives 95 % confidence intervals of $294.07 < \phi < 294.76$ µm for the diameter and $7.551 < d < 7.667$ µm for the depth.

These combined uncertainty distributions are approximately Gaussian and adding the uncertainties in quadrature gives a standard deviation of 163 nm in diameter and 26 nm in depth, which is in agreement with the Monte Carlo approach and suggests that the law of propagation of uncertainty is a good approximation in this case.

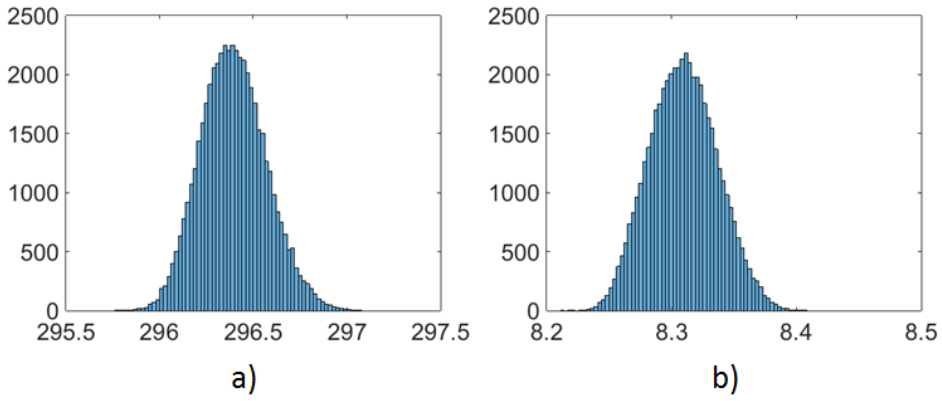

Figure 6.15: Histogram of Monte Carlo simulation results for a) diameter and b) depth. x axis in µm*. y axis is in counts.*

As well as the uncertainty, the combined bias should also be estimated. By taking the sum of individual biases, the total bias can be calculated as −372 nm in diameter and −72 nm in depth for the dimple shown in figure 6.15.

Table 6.2 summarises the uncertainty and bias contributions for all three dimples used. There are differences between the three dimples, primarily in the uncertainty and bias of the diameter. These differences are believed to be primarily due to the variation in topography between the dimples, which will affect the sensitivity of the segmentation

to the changes caused by the metrological characteristics and, therefore, the uncertainty. Based on the differences between dimples it is recommended, as in chapter 5, to take maximum uncertainty estimate as the estimate of the uncertainty of the surface in order to provide a conservative uncertainty estimate.

As well as the uncertainty, the bias varies between the three dimples measured. Again, this variation is believed to be due to how the changing topography interacts with the segmentation method and metrological characteristics. Given the instability in the bias, there should be an additional uncertainty contribution from the uncertainty in the value of the bias when any other surface is measured. This uncertainty could be modelled in many ways. Here a curvilinear trapezoid has been used, which is similar to a uniform distribution but better accounts for the uncertainty in the limits due to the small number of reference points Giusca et al. 2012a; JCGM 2008b. The curvilinear trapezoid, Ctrap$(Max(bias), Min(bias), \sigma_{bias})$ has standard deviation

$$\sigma_{Ctrap} = \sqrt{\frac{W^2}{12} + \frac{\sigma_{bias}^2}{9}} \tag{6.3}$$

where W is the nominal width of the underlying rectangular distribution given by $W = Max(bias) - Min(bias)$ and σ_{bias} is the standard deviation of the combined bias mea-

Table 6.2: Summary of results showing standard uncertainties and bias due to each characteristic as well as the combined uncertainties and biases.

	σ(D)			σ(depth)			bias(D)			bias(depth)		
Dimple	1	2	3	1	2	3	1	2	3	1	2	3
Noise (nm)	74	90	64	2	2	2	80	154	88	12	12	12
Flatness (nm)	95	133	113	17	16	16	-416	-395	-488	-76	-77	-79
Vert. scales (nm)	13	29	48	20	22	23	15	<1	-45	-8	-9	-10
Lat. scales (nm)	64	65	75	<1	<1	<1	-51	-68	-58	<1	<1	<1
Filter effect (nm)	13	14	18	<1	<1	<1	-	-	-	-	-	-
Repeatability of lat. scales (nm)	58	58	56	<1	<1	<1	-	-	-	-	-	-
Uncertainty in bias (nm)	65	65	65	2	2	2	-	-	-	-	-	-
Combined (nm)	163	196	181	26	27	28	-372	-309	-503	-72	-74	-77
Combined with correlation (nm)	197	236	224	27	27	28	-	-	-	-	-	-

surements. The combined biases give $W = 194$ nm and $\sigma_{bias} = 99$ nm in diameter, and $W = 5$ nm and $\sigma_{bias} = 2.5$ nm in depth. These values result in an additional uncertainty of 65 nm in diameter and 2 nm in depth, which is included in the uncertainty in table 6.2. Correction of any measurements on this surface could then be performed using the mean bias values of -395 nm in diameter and -74 nm in depth.

The results in table 6.2 do not include the effects of the lateral scale amplification factor. With the available data, the best estimate of 1 % uncertainty in the lateral amplification factor would far outweigh any other contribution. However, as the amplification factor is a constant of the measurement system it does not affect the reproducibility of the system and will not decrease when taking uncertainty of the mean from multiple measurements. Therefore, the lateral amplification factor is only directly important when measuring the absolute values of a feature, not when comparing between different features measured with the same instrument. There may still be indirect effects from the uncertainty in the amplification factor, which do vary between features due to the dependence on local topography, such as the filter effects, but these have been considered separately and are included in table 6.2. A similar argument could also be made for the vertical scale amplification factor. However, the indirect effects in the vertical scales are much stronger, due to the complex interaction between the height values and segmentation algorithm. Therefore, further research into how the vertical scale amplification changes between measurements may be of interest.

6.5 Correlation

The above analysis assumes that all the metrological characteristics are independent of each other. However, due to their position dependence it is suspected that there may be some correlation between the effect of flatness and lateral scales.

This effect can be investigated by using the same set of random positions to model the effect of both flatness deviation and the lateral scales. Then combining the uncertainty using a Monte Carlo approach, drawing the same sample for both the flatness and lateral scales. In this way, the contributions due to flatness and lateral scales will always be evaluated at the same position and so will be correlated. Table 6.2 also shows the combined uncertainties once this correlation has been accounted for. As can be seen, the diameter has significantly higher uncertainty once correlation is added, which suggests that the metrological characteristics cannot be viewed as independent in this case. The effect of correlation on the depth is negligible, which is unsurprising considering that the effect of

the lateral scales on depth is almost zero.

Due to the dependence on the lateral scales, it is expected that the effect of filtering is also correlated in a similar way. However, as the effect of filtering is not dependant on position, it is not straightforward to consider this correlation. In any case, the effect of filtering is small and, therefore, the changes due to any correlation with position are likely to be negligible.

6.6 Comparison with reproducibility

Once correlation is accounted for, the approach in this chapter can be compared with the analysis in chapter 5. Using metrological characteristics gives a standard uncertainty of 222 nm in diameter and 28 nm in depth, compared to 202 nm in diameter and 18 nm in depth from the repeatability tests on the confocal microscope.

These results are approximately equivalent with metrological characteristics giving a slightly higher uncertainty, which is expected as the vertical scales and uncertainty in the bias are not considered in the reproducibility and other effects may be underaccounted for.

Additionally, the metrological characteristics provide an estimate of the bias in the measurements, which reproducibility measurements cannot give. It should also be noted that the changing topography alters the bias values between measurements and contributes an additional uncertainty term. For these measurements, this additional uncertainty did not significantly effect the combined uncertainty but for other cases it may be more important.

6.7 Discussion

The procedure described in this chapter could be applied to any other structured surface to determine uncertainty in geometric properties of those surfaces. The uncertainty is estimated using the metrological characteristics of the instrument, which are mostly unaffected by the sample used and are the same measurements that would be required to assess uncertainty in non-structured surfaces. Therefore, the assessment of metrological characteristics should already be part of the instrument calibration process and require minimal extra work. The main difference between the method used here and that for non-structured surfaces is in how the metrological characteristics are propagated into the final uncertainty.

The main drawback of the approach presented in this chapter is that it relies heavily on Monte Carlo techniques to propagate the uncertainties. While this is necessary, due to non-differentiable nature of segmentation algorithms, Monte Carlo techniques are also very computationally expensive which makes the approach used here somewhat time consuming to implement in practise. For this reason, one interesting line of future research would be to investigate possible models and approximations that could avoid or reduce the need for Monte Carlo simulations, such as using a linear approximation over a few positions.

In this chapter, the microscope was not corrected based on the measured metrological characteristics. Firstly, the associated uncertainty, even with the uncorrected metrological characteristics, is still small compared to the variation in the geometric properties across the surface, which chapter 5 shows is on the order of 4 µm in the diameter. The main aim of this chapter is to demonstrate the principle of calculating the uncertainty in structured surfaces, not to provide the lowest uncertainty possible. Additionally, for some uncertainty components, particularly the lateral distortion, correcting the error is quite difficult and introduces many more complications in the calculation stage, as the grid is no longer uniform and must be considered as a point cloud.

While the magnitudes of the metrological characteristics are quite large, compared to values for some other instruments, they are mostly not unexpected for this type of scanning confocal microscope. The one surprising result is the non-linearity in the z scales. The 500 nm, 1.2 µm and 2.4 µm steps gave significantly worse results than the other three steps, with over 70 nm residual, which is very high. This error was determined to be a material effect and dependant on the instrument brightness used. Further research would be necessary to identify the base cause of this effect and how it effects real measurement results.

It should also be remembered that the use of metrological characteristics for surface metrology is still a relatively new idea and relies on certain assumptions about the measurement set-up. The main issue of concern is that some of the metrological characteristics are not measurement independent. This means that characteristics such as noise and flatness will change depending on the measurement set-up used. Such changes can include many factors, such as changes in illumination and material between artefacts and real measurement samples. Ideally, measurements would be performed under identical conditions when calibrating and preforming measurements. However, in reality this is never possible and there is often a desire to perform as general a calibration as possible to avoid having to repeat it for each measurement. In theory, these effects could be in-

corporated into the uncertainty in the metrological characteristics. However, significant further research would be required to determine how various factors effect the metrological characteristics and such effects are likely to be somewhat instrument dependant as well.

Finally, throughout this chapter the choice of segmentation algorithm and associated choices, such as filtering and height threshold, have been considered fixed. These choices will have a strong influence on the geometric properties of the features and, under the approach used here, the segmentation method defines the position of the feature boundary. An alternative viewpoint is possible, where there is a 'true' feature which the segmentation method aims to determine. Under this approach, there may be some uncertainty in the correct threshold, filter and other algorithmic choices to best identify the feature, resulting in additional uncertainty in the geometric properties of the feature. Such an approach is not considered in this chapter as it is not clear how to define the ground truth or the uncertainties associated with the thresholds, etc. Further knowledge of the functional performance of the surface may help in determining these values.

6.8 Conclusion

This chapter has presented an approach to determine the uncertainty in the measurement of geometric properties of structured surfaces using the metrological characteristics of the measuring instrument. The main conclusions are as follows:

1. Four metrological characteristics: measurement noise, flatness deviation, vertical and lateral scale deviations were measured using standard techniques and then propagated into the uncertainty in the geometric properties of a feature using novel, Monte Carlo based methods.

2. The single values normally reported for metrological characteristics, such as Sq of the noise and Sz of the flatness, where not sufficient to propagate into the uncertainties and additional information such as segmenting the noise map and using the full flatness and lateral distortion maps was necessary in order to estimate the uncertainty.

3. The use of metrological characteristics also allows for estimation of measurement bias by comparing the results with those where the effect of the metrological characteristic has been corrected, which is not possible when performing a reproducibility

assessment. This bias was found to vary between dimples due to the interaction with the local topography. Therefore, to allow for measurement of other dimples on the same surface, the uncertainty in the bias was also added to the overall uncertainty.

4. Correlation between the flatness deviation and lateral deviation was determined by modelling these effects using the same set of positions and drawing from the joint distribution when combining the uncertainties. It was found that there was significant correlation as both these effects are highly dependant on the position of the feature in the field of view. This is an important result as it shows that metrological characteristics cannot always be viewed as independent from each other.

5. The combined uncertainty of all these effects was calculated using both Monte Carlo techniques and the standard law of propagation of uncertainty, although the law of propagation of uncertainty cannot easily account for correlation between components. The combined uncertainty was 229 nm in diameter and 28 nm in depth. These values are slightly higher than the results from the reproducibility analysis in chapter 5, due to the additional effects of vertical scales and other factors which cannot easily be fully determined in the reproducibility study.

6. The amplification of the lateral scales was not measured directly, as a suitably calibrated grid was not available, and is not included in the combined uncertainty. The manufacter's calibration of the lateral scales, which gives a maximum error of 1 %, could be used as an estimate of the uncertainty in the amplification of the lateral scales. However, this value is probably an overestimate as it will include some non-linear effects which are already accounted for in the lateral deviation calibration. If this value were used, it would dominate the uncertainty in the diameter. However, amplification does not contribute to the repeatability of a measurement and, therefore, has no direct effect when comparing between features with the same instrument, only when an absolute value is needed. Some indirect effects of the lateral amplification, such as filter effects may vary between different measurements, but these are minor effects and are already included in the combined uncertainty.

Chapter 7

Discussion, conclusion and future work

This thesis has successfully investigated the effects that influence the geometric properties of structured surfaces. Additionally, two methods have been developed to assess the uncertainty in these geometric properties. This chapter highlights the main conclusions and contributions and discusses possible areas of future research.

7.1 Conclusions

The main conclusions of this thesis are as follows:

1. A review of the current literature with respect to characterisation of structured surfaces was performed. It concluded that:

   (a) Field parameters are generally viewed as unsuitable for the characterisation of structured surfaces as they do not differentiate between surface features and background surface.

   (b) Identification of individual features and calculation of geometric parameters is the most common method to characterise structured surfaces in the current literature. The most common segmentation methods were discussed in more detail.

   (c) Pattern matching techniques are another approach which can successfully characterise structured surfaces. However, this approach requires a CAD model or similar to compare the surface to.

2. The effects of control parameters, such as threshold level and filtering, on the geometric properties of a feature were considered. These effects have not previously been explicitly considered for structured surfaces. It was found that:

   (a) The effect of changing thresholds and similar algorithm control parameters was investigated. For height and gradient thresholding methods, changing the threshold value had significant effect on the feature size. On the other hand, changing smoothing parameters for the active contour method had little effect, as this method always seeks to minimise the variance between regions. For morphological segmentation, provided that the feature is still segmented successfully, the merging parameters can vary significantly without effecting the feature boundary as the position of ridges is unaffected.

   (b) The effect of filtering was also estimated using a range of Gaussian filters. The filter cut-off was found to have significant effect on the feature size. For the gradient threshold, the dimple diameter changed by approximately 20 μm over the range of cut-offs used. However, for the other methods the effect was smaller, approximately 1 or 2 μm change. Most notably, there was no simple relationship between filter size and measured geometric properties; the interaction between local topography and the filter can produce a larger dimple with a smaller filter or vice versa.

3. As well as the effects of instrument control parameters, the measurement uncertainty in characterising structured surfaces was estimated, which has not previously been considered. An initial assessment of the measurement uncertainty was made by a reproducibility analysis, where a dimple was repeatedly measured over a 5×5 grid.

   (a) The results of the reproducibility tests give an expanded uncertainty ($k = 2$) of 180 nm in diameter and 2.5 μm for the out of roundness using the 300 μm disk. 50 and 150 μm diameter disks were also assessed with similar results.

   (b) Manufacturing reproducibility was also estimated by measuring a sample of 100 dimples and was found to be much larger than the measurement reproducibility, approximately 5 to 10 μm, depending on dimple size. Therefore, the measurement uncertainty does not significantly affect the estimation of manufacturing reproducibility in this case. However, it is still important to estimate uncertainty as it may be more important for other surfaces.

(c) The position dependence of the geometric properties was also considered. While in many cases the correction was similar across the different dimples, it was also less than the variance at a single position. Additionally, no absolute reference was available to determine a suitable value to correct the properties to. Therefore, no attempt to actually correct the data using the position dependence was made.

(d) The effect of varying height and gradient thresholds on the uncertainty was also assessed. For gradient thresholds, there was a region between gradient 0.2 and 0.3 where the uncertainty was minimised. Outside this region, the uncertainty varied significantly between dimples and was generally unstable. For the height threshold, the uncertainty was more stable. However, there was slightly lower uncertainty in the range -1 to -0.5 μm.

(e) Several drawbacks to estimating uncertainty by a reproducibility study were identified. Firstly, some uncertainty components will not be accounted for using the reproducibility tests. Specifically, those that do not change significantly between repeated measurements and are not position dependant, such as vertical scale effects. If these sources are believed to be significant, another method would be needed to assess them. Secondly, the reproducibility cannot assess any bias in the measurements and, therefore, systematic measurement errors cannot easily be accounted for.

4. To deal with the drawbacks of using the reproducibility, an alternative novel approach to assess uncertainty in structured surfaces using metrological characteristics was considered. This approach provides a more comprehensive assessment of the uncertainty, accounting for additional uncertainty components that are not included in the reproducibility analysis and allowing for assessment of bias in the measurements.

(a) Four metrological characteristics: measurement noise, flatness deviation, vertical scales and lateral scales were measured using standard techniques.

(b) Monte Carlo and other techniques were used to propagate these metrological characteristics into uncertainties in the geometric properties using three example dimples. In order to account for the nature of structured surfaces, more information than the single values normally reported for metrological characteristics was used, such as using the full flatness map and segmenting the noise

map into multiple regions.

(c) By correcting for the effect of the metrological characteristics, the bias introduced to the measurements could be calculated. There was significant variation between the bias of the three measured dimples due to variations in the local topography and the interaction with the metrological characteristics and segmentation algorithm. The average bias was -395 nm in diameter and -74 nm in depth and the variation in bias provided an additional contribution to the uncertainty.

(d) Uncertainty due to each individual metrological characteristic was combined to give a combined uncertainty of 229 nm in diameter and 28 nm in depth. The main contributor to uncertainty was the flatness deviation.

(e) The amplification in the lateral scales was not directly assessed as no calibrated grid was available and is not included in the combined uncertainty. The manufacturer's calibrated value could be used, but the value of 1 % is likely an overestimate as it includes some non-linearity components which are already measured. This value is far larger than any other uncertainty components in the diameter, but does not effect the reproducibility of the instrument and is only significant when absolute values are needed.

(f) The results calculated using metrological characteristics were compared with the reproducibility results. The results using the two methods are similar and as expected the metrological characteristics give a slightly higher uncertainty due to the inclusion of the effect of vertical scales and uncertainty in bias, which are not considered in the reproducibility analysis.

7.2 Future work

This thesis has advanced the state of the art in characterisation of structured surfaces. However, there are many possible avenues for further research in this area. These include:

- There are many advanced segmentation and characterisation strategies present in the image processing and computer vision literature that have not yet been applied to surface metrology. These techniques may allow for more relevant characterisation with fewer problems such as choice of threshold. In particular, support vector machines (SVM) (Yu et al. 2011b; Song and Civco 2004), neural networks (Cheng

et al. 1996) and fuzzy logic (Yang and Huang 2012; Mohamed et al. 1999; Li et al. 1994) appear to be promising approaches.

- The effect of choice of thresholds and filter size on geometric properties of structured surfaces has been investigated. However, these results are limited to a particular test case and cannot easily be generalised to different surface structures as the effects are dependant on the local topography. Further research may allow for a more general model of the effect of these parameters to be developed. A modelling based approach would likely be necessary due to the vast range of possible structures, which make it logistically impossible to manufacture and measure a complete range of structures.

- When considering uncertainty using metrological characteristics, extensive use was made of Monte Carlo simulations. This makes the analysis very computationally intensive and time consuming. Further work to develop a simplified model to propagate the uncertainty would allow for much simpler and faster analysis of uncertainty than is possible with the techniques presented here. One possibility would be to the effects of metrological characteristics, such as flatness, are evaluated at a fixed number of positions and the maximum and minimum values are used to estimate the uncertainty. Such an approach would rely on having a good model of how the metrological characteristics affect the results, in order to develop a reasonable model for the uncertainty distribution. Additionally, non-linear effects which this thesis have shown can be significant would have to be considered as they could invalidate a linear model.

- The effect of the lateral resolution on measurements and their uncertainty has not been considered. While the lateral resolution can be measured (Giusca and Leach 2013), it is difficult to relate to distortion of the surface, as it does not provide full information about the measurement system (Foreman et al. 2013). Measurement of the instrument transfer function could provide more complete information. However, such measurements are challenging and do not account for reconstruction errors and other non-linear effects, that can cause significant errors (Gao et al. 2008; Coupland and Lobera 2010). Further fundamental research is needed to better understand these non-linear effects. Many commercial instruments currently employ heuristic corrections to eliminate many of these effects. However, these corrections cannot always be effective, see for example the step height measurements in section 6.2.3,

and often make it hard for the end user to understand what the problems with their measurements really are. That said, these corrections are also necessary to provide the fast, high quality measurements required by industry and so a balance must be struck between fundamental research to better characterise the instruments and the current needs of industry.

7.3 Publications

The results of this thesis have resulted in and contributed to several publications as follows:

- G D MacAulay, N Senin, C L Giusca, R K Leach, A Ivanov (2015) Review of feature boundary identification techniques for the characterization of tessellated surfaces, Surf. Topog.: Met. Prop., **3**, 013002
- G D MacAulay, N Senin, C L Giusca, R K Leach (2014) Comparison of segmentation techniques to determine the geometric parameters of structured surfaces, Surf. Topog.: Met. Prop., **2**, 044004
- G D MacAulay, N Senin, C L Giusca, R K Leach (2015) Investigation of micro-manufacture process performance via innovative surface characterisation methods, Measurement, *under review*
- G D MacAulay, C L Giusca (2016) Assessment of uncertainty in structured surfaces using metrological characteristics, *to be submitted to CIRP Annals 2016*
- N Senin, G D MacAulay, C L Giusca, R K Leach (2014) On the characterisation of periodic patterns in tessellated surfaces, Surf. Topog.: Met. Prop., **2**, 025005

Bibliography

Alicona Infinite Focus. URL: http://www.alicona.co.uk/home/products/infinitefocus.html (visited on 11/06/2015).

Anchini, R., C. Liguori and A. Paolillo (2007). "Evaluation of the Uncertainty of Edge-Detector Algorithms". In: *IEEE Transactions on Instrumentation and Measurement* 56.3, pp. 681–688. DOI: 10.1109/TIM.2007.894909.

Andersson, P., J. Koskinen, S. Varjus, Y. Gerbig, H. Haefke, S. Georgiou, B. Zhmud and W. Buss (2007). "Microlubrication effect by laser-textured steel surfaces". In: *Wear* 262.3-4, pp. 369–379. DOI: 10.1016/j.wear.2006.06.003.

Artigas, R. (2011). "Imaging Confocal Microscopy". In: *Optical Measurement of Surface Topography*. Ed. by R. K. Leach. Springer. Chap. 11.

Barre, F. and J. Lopez (2000). "Watershed lines and catchment basins: a new 3D-motif method". In: *International Journal of Machine Tools and Manufacture* 40.8, pp. 1171–1184. DOI: 10.1016/S0890-6955(99)00118-2.

Bechert, D. W., M. Bruse and W. Hage (2000). "Experiments with three-dimensional riblets as an idealized model of shark skin". In: *Experiments in Fluids* 28.5, pp. 403–412. DOI: 10.1007/s003480050400.

Bergholm, F. (1987). "Edge Focusing". In: *IEEE transactions on pattern analysis and machine intelligence* 9.6, pp. 726–41.

Besl, P. J. and H. D. McKay (1992). "A Method for Registration of 3-D Shapes". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.2, pp. 239–256. DOI: 10.1109/34.121791.

Bieniek, A and A Moga (2000). "An efficient watershed algorithm based on connected components". In: *Pattern Recognition* 33, pp. 907–916.

Blateyron, F. (2011). "Chromatic Confocal Microscopy". In: *Optical Measurement of Surface Topography*. Ed. by R. K. Leach. Springer. Chap. 5.

— (2013a). "The Areal Feature Parameters". In: *Characterisation of Areal Surface Texture*. Ed. by R. K. Leach. Springer. Chap. 3, pp. 45–66.

— (2013b). “The Areal Field Parameters”. In: *Characterisation of Areal Surface Texture*. Ed. by R. K. Leach. Springer. Chap. 2, pp. 15–44.

Bleau, A. and J. L. Leon (2000). “Watershed-Based Segmentation and Region Merging”. In: *Computer Vision and Image Understanding* 77.3, pp. 317–370. DOI: 10.1006/cviu.1999.0822.

Blunt, L. and P. J. Scott (2012). “Characterisation of structured geometry surface features on micro and nano scale surfaces”. In: *International Journal of Nanomanufacturing* 8.5/6, p. 359. DOI: 10.1504/IJNM.2012.051111.

Blunt, L. and S. Xiao (2011). “The use of surface segmentation methods to characterise laser zone surface structure on hard disc drives”. In: *Wear* 271.3-4, pp. 604–609. DOI: 10.1016/j.wear.2010.06.014.

Bovik, A. (2010). *Handbook of Image and Video Processing*. 2nd. Academic Press.

Brinkmann, S, H Bodschwinna and H. W Lemke (2001). “Accessing roughness in three-dimensions using Gaussian regression filtering”. In: *International Journal of Machine Tools and Manufacture* 41.13-14, pp. 2153–2161. DOI: 10.1016/S0890-6955(01)00082-7.

Brinksmeier, E., R. Gläbe and C. Flucke (2008). “Manufacturing of molds for replication of micro cube corner retroreflectors”. In: *Production Engineering* 2.1, pp. 33–38. DOI: 10.1007/s11740-008-0082-8.

Bruzzone, A. A. G., H. L. Costa, P. M. Lonardo and D. A. Lucca (2008). “Advances in engineered surfaces for functional performance”. In: *CIRP Annals - Manufacturing Technology* 57.2, pp. 750–769. DOI: 10.1016/j.cirp.2008.09.003.

Canny, J. (1986). “A Computational Approach to Edge Detection”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.6, pp. 679–698. DOI: 10.1109/TPAMI.1986.4767851.

Caselles, V., R. Kimmel and G. Sapiro (2001). “Geodesic Active Contours”. In: *IEEE International Conference on Computer Vision*. Vol. 10. 10. IEEE Comput. Soc. Press, pp. 694–699. DOI: 10.1109/ICCV.1995.466871.

Chan, T. F. and L. A. Vese (2001). “Active Contours Without Edges.” In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 10.2, pp. 266–77. DOI: 10.1109/83.902291.

Cheng, K.-S., J.-S. Lin and C.-W. Mao (1996). “The application of competitive Hopfield neural network to medical image segmentation”. In: *IEEE Transactions on Medical Imaging* 15.4, pp. 560–567. DOI: 10.1109/42.511759.

Chua, C. S. and R. Jarivs (1997). “Point Signatures : A New Representation for 3D Object Recognition”. In: *International Journal of Computer Vision* 25.1, pp. 63–85. DOI: 10.1023/A:1007981719186.

Cochran, W. G. (1934). “The distribution of quadratic forms in a normal system, with applications to the analysis of covariance”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 30.02, p. 178. DOI: 10.1017/S0305004100016595.

Coupland, J. M. and J. Lobera (2010). “Measurement of Steep Surfaces Using White Light Interferometry”. In: *Strain* 46.1, pp. 69–78. DOI: 10.1111/j.1475-1305.2008.00595.x.

Crouch, A. S., D. Miller, K. J. Luebke and W. Hu (2009). “Correlation of anisotropic cell behaviors with topographic aspect ratio”. In: *Biomaterials* 30.8, pp. 1560–1567. DOI: 10.1016/j.biomaterials.2008.11.041.

Danzl, R., F. Helmli and S. Scherer (2011). “Focus variation - a Robust Technology for High Resolution Optical 3D Surface Metrology”. In: *Strojniski Vestnik/Journal of Mechanical Engineering* 57.3, pp. 245–256. DOI: 10.5545/sv-jme.2010.175.

De Santo, M., C. Liguori and A. Pietrosanto (2000). “Uncertainty Characterization in Image-Based Measurements: A Preliminary Discussion”. In: *IEEE Transactions on Instrumentation and Measurement* 49.5, pp. 1101–1107. DOI: 10.1109/19.872937.

De Santo, M., C. Liguori, A. Paolillo and A. Pietrosanto (2004). “Standard uncertainty evaluation in image-based measurements”. In: *Measurement* 36.3-4, pp. 347–358. DOI: 10.1016/j.measurement.2004.09.011.

Dietrich, C. F. (1991). *Uncertainty, Calibration and Probability: The Statistics of Scientific and Industrial Measurement*. 2nd. CRC Press.

Dietzsch, M., K. Papenfub and T. Hartmann (1998). *The MOTIF-method - a suitable description for functional, manufactural and metrological requirements*. DOI: 10.1016/S0890-6955(97)00110-7.

Downs, M. J., A. B. Forbes and J. E. Siddle (1999). “The verification of a high-precision two-dimensional position measurement system”. In: *Measurement Science and Technology* 9.7, pp. 1111–1114. DOI: 10.1088/0957-0233/9/7/018.

Etsion, I and L Burstein (1996). “A model for mechanical seals with regular microsurface structure”. In: *Tribology Transactions* 39.3, pp. 677–683. DOI: 10.1080/10402009608983582.

Etsion, I, Y Kligerman and G Halperin (1999). “Analytical and Experimental Investigation of Laser-Textured Mechanical Seal Faces”. In: *Tribology Transactions* 42.3, pp. 511–516. DOI: 10.1080/10402009908982248.

Etsion, I. (2005). “State of the Art in Laser Surface Texturing”. In: *Journal of Tribology* 127.1, pp. 248–253. DOI: 10.1115/1.1828070.

— (2013). “Modeling of surface texturing in hydrodynamic lubrication”. In: *Friction* 1.3, pp. 195–209. DOI: 10.1007/s40544-013-0018-y.

Evans, C. J. and J. B. Bryan (1999). “Structured, Textured or Engineered surfaces”. In: *CIRP Annals - Manufacturing Technology* 48.2, pp. 541–556. DOI: 10.1016/S0007-8506(07)63233-8.

Extrand, C. W., S. I. Moon, P Hall and D Schmidt (2007). “Superwetting of Structured Surfaces”. In: *Langmuir : the ACS journal of surfaces and colloids* 23.17, pp. 8882–90. DOI: 10.1021/la700816n.

Foreman, M. R., C. L. Giusca, J. M. Coupland, P. Török and R. K. Leach (2013). “Determination of the transfer function for optical surface topography measuring instruments - a review”. In: *Measurement Science and Technology* 24.5, p. 052001. DOI: 10.1088/0957-0233/24/5/052001.

Fowell, M. T., S. Medina, A. V. Olver, H. A. Spikes and I. G. Pegg (2012). “Parametric study of texturing in convergent bearings”. In: *Tribology International* 52, pp. 7–16. DOI: 10.1016/j.triboint.2012.02.013.

Gao, F, R. K. Leach, J Petzing and J. M. Coupland (2008). “Surface measurement errors using commercial scanning white light interferometers”. In: *Measurement Science and Technology* 19.1, p. 015303. DOI: 10.1088/0957-0233/19/1/015303.

Gao, H., P. Xue and W. Lin (2004). “A new marker-based watershed algorithm”. In: *IEEE International Symposium on Circuits and Systems*. IEEE, pp. II–81–4. DOI: 10.1109/ISCAS.2004.1329213.

Gauch, J. M. (1999). “Image Segmentation and Analysis via Multiscale Gradient Watershed Hierarchies.” In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 8.1, pp. 69–79. DOI: 10.1109/83.736688.

Gerig, G., O. Kubler, R. Kikinis and F. A. Jolesz (1992). “Nonlinear anisotropic filtering of MRI data”. In: *IEEE transactions on medical imaging* 11.2, pp. 221–32. DOI: 10.1109/42.141646.

Geringer, J., P. Arnecke and J. Seewig (2013). “Novel parameters to assess feature size on functional surfaces”. In: *Surface Topography: Metrology and Properties* 1.1, p. 015004. DOI: 10.1088/2051-672X/1/1/015004.

Giusca, C. L. and R. K. Leach (2012). *NPL Good Practice Guide 128: Calibration of the metrological characteristics of Imaging Confocal Microscopes*.

Giusca, C. L. and R. K. Leach (2013). "Calibration of the scales of areal surface topography measuring instruments: part 3. Resolution". In: *Measurement Science and Technology* 24.10, p. 105010. DOI: 10.1088/0957-0233/24/10/105010.

Giusca, C. L., R. K. Leach, F. Helary, T. Gutauskas and L. Nimishakavi (2012a). "Calibration of the scales of areal surface topography-measuring instruments: part 1. Measurement noise and residual flatness". In: *Measurement Science and Technology* 23.3, p. 035008. DOI: 10.1088/0957-0233/23/3/035008.

Giusca, C. L., R. K. Leach and F. Helery (2012b). "Calibration of the scales of areal surface topography measuring instruments: part 2. Amplification, linearity and squareness". In: *Measurement Science and Technology* 23.6, p. 065005. DOI: 10.1088/0957-0233/23/6/065005.

Giusca, C. L., J. D. Claverley, W. Sun, R. K. Leach, F. Helmli and M. P. J. Chavigner (2014). "Practical estimation of measurement noise and flatness deviation on focus variation microscopes". In: *CIRP Annals - Manufacturing Technology* 63.1, pp. 545–548. DOI: 10.1016/j.cirp.2014.03.086.

Goldenberg, R., R. Kimmel, E. Rivlin and M. Rudzsky (2001). "Fast Geodesic Active Contours." In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 10.10, pp. 1467–75. DOI: 10.1109/83.951533.

Gonzalez, R. C. and R. E. Woods (2008). *Digital Image Processing*. 3rd. Prentice Hall.

Greiner, C., A. del Campo and E. Arzt (2007). "Adhesion of Bioinspired Micropatterned Surfaces: Effects of Pillar Radius, Aspect Ratio, and Preload." In: *Langmuir* 23.7, pp. 3495–502. DOI: 10.1021/la0633987.

Groot, P. de (2011). "Coherence Scanning Interferometry". In: *Optical Measurement of Surface Topography*. Ed. by R. K. Leach. Springer. Chap. 9.

Grubbs, F. E. (1950). "Sample Criteria for Testing Outlying Observations". In: *The Annals of Mathematical Statistics* 21.1, pp. 27–58. DOI: 10.1214/aoms/1177729885.

Haitjema, H (1998). "Uncertainty analysis of roughness standard calibration using stylus instruments". In: *Precision Engineering* 22.97, pp. 110–119. DOI: 10.1016/S0141-6359(97)00090-1.

Haitjema, H. (2013). "Uncertainty estimation of 2.5-D roughness parameters obtained by mechanical probing". In: *International Journal of Precision Technology* 3.4, p. 403. DOI: 10.1504/IJPTECH.2013.058260.

Harasaki, A. and J. C. Wyant (2000). "Fringe modulation skewing effect in white-light vertical scanning interferometry". In: *Applied Optics* 39.13, p. 2101. DOI: 10.1364/AO.39.002101.

Harris, P., R. K. Leach and C. L. Giusca (2010). *NPL Report MS 8 - Uncertainty evaluation for the calculation of a surface texture parameter in the profile case*. Tech. rep.

Hartmann, W. and A. Loderer (2014). "Automated extraction and assessment of functional features of areal measured microstructures using a segmentation-based evaluation method". In: *Surface Topography: Metrology and Properties* 2.4, p. 044001. DOI: 10.1088/2051-672X/2/4/044001.

Hartmann, W. and A. Loderer (2013). "Measurement Procedures and Evaluations Methods for the Extraction and Assessment of Functional Features of Microstructured Surfaces". In: *11th International Symposium on Measurement and Quality Control*, p. 61.

He, B., W. Chen and Q. Jane Wang (2008). "Surface Texture Effect on Friction of a Microtextured Poly(dimethylsiloxane) (PDMS)". In: *Tribology Letters* 31.3, pp. 187–197. DOI: 10.1007/s11249-008-9351-0.

Helmli, F. (2011). "Focus Variation Instruments". In: *Optical Measurement of Surface Topography*. Ed. by R. K. Leach. Springer. Chap. 7.

Henning, A., C. Giusca, A. Forbes, I. Smith, R. Leach, J. Coupland and R. Mandal (2013). "Correction for lateral distortion in coherence scanning interferometry". In: *CIRP Annals - Manufacturing Technology* 62.1, pp. 547–550. DOI: 10.1016/j.cirp.2013.03.026.

Hodge, V. and J. Austin (2004). "A Survey of Outlier Detection Methodologies". In: *Artificial Intelligence Review* 22.2, pp. 85–126. DOI: 10.1023/B:AIRE.0000045502.10941.a9.

Huang, Y.-F. et al. (2007). "Improved broadband and quasi-omnidirectional anti-reflection properties with biomimetic silicon nanostructures." In: *Nature nanotechnology* 2.12, pp. 770–774. DOI: 10.1038/nnano.2007.389.

Hughes, I. and T. P. A. Hase (2010). *Measurements and their uncertainties: a practical guide to modern error analysis*. Oxford University Press.

ISO (1997). *ISO 4287 Geometrical Product Specifications (GPS) – Surface texture: Profile method – Terms, definitions and surface texture parameters*.

— (2012). *ISO 25178-2 Geometrical Product Specifications (GPS) - Surface Texture: Areal - Part 2: Terms, definintions and surface texture parameters*.

— (2013). *ISO 1101 Geometric Product Specifications (GPS) - Geometrical tolerancing - Tolerances of form, orientation, location and run-out*.

ISO (2015). *ISO CD 25178-600: Geometrical product specifications (GPS) - Surface texture: Areal - part 600: Metrological charcteristics for areal topography measuring methods.*

Ismail, M. F., K. Yanagi and A. Fujii (2010). "An outlier correction procedure and its application to areal surface data measured by optical instruments". In: *Measurement Science and Technology* 21.10, p. 105105. DOI: 10.1088/0957-0233/21/10/105105.

JCGM (2008a). *Guide to the expression of uncertainty in measurement.*

— (2008b). *Supplement 1 to the Guide to the expression of uncertainty in measurement - Propagation of distributions using a Monte Carlo method.*

— (2012). *International vocabulary of metrology - Basic and general concepts and associated terms.*

— (2015). *Evaluation of measurement data - Guide to uncertainty in measurement (Committee Draft).*

Jackway, P. T. (1996). "Gradient Watersheds in Morphological Scale-Space". In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 5.6, pp. 913–21. DOI: 10.1109/83.503908.

Jiang, X. J. and D. J. Whitehouse (2012). "Technological shifts in surface metrology". In: *CIRP Annals - Manufacturing Technology* 61.2, pp. 815–836. DOI: 10.1016/j.cirp.2012.05.009.

Jiang, X., P. J. Scott, D. J. Whitehouse and L. Blunt (2007). "Paradigm shifts in surface metrology. Part II. The current shift". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 463.2085, pp. 2071–2099. DOI: 10.1098/rspa.2007.1873.

Jiang, X., X. Zhang and P. J. Scott (2010). "Template matching of freeform surfaces based on orthogonal distance fitting for precision metrology". In: *Measurement Science and Technology* 21.4, p. 045101. DOI: 10.1088/0957-0233/21/4/045101.

Kass, M., A. Witkin and D. Terzopoulos (1988). "Snakes: Active Contour Models". In: *International Journal of Computer Vision* 1.4, pp. 321–331. DOI: 10.1007/BF00133570.

Kiyono, S., W. Gao, S. Zhang and T. Aramaki (2000). "Self-calibration of a scanning white light interference microscope". In: *Optical Engineering* 39.10, p. 2720. DOI: 10.1117/1.1290471.

Kong, L. B., C. F. Cheung, X. Q. Jiang, W. B. Lee, S. To, L. Blunt and P. Scott (2010). "Characterization of surface generation of optical microstructures using a pattern and feature parametric analysis method". In: *Precision Engineering* 34.4, pp. 755–766. DOI: 10.1016/j.precisioneng.2010.04.005.

Le Goic, G., C. A. Brown, H. Favreliere, S. Samper and F. Formosa (2013). "Outlier filtering: a new method for improving the quality of surface measurements". In: *Measurement Science and Technology* 24.1, p. 015001. DOI: 10.1088/0957-0233/24/1/015001.

Leach, R. K., C. L. Giusca, H. Haitjema, C. Evans and X. Jiang (2015). "Calibration and verification of areal surface texture measuring instruments". In: *CIRP Annals - Manufacturing Technology* 64.2.

Leach, R., ed. (2011). *Optical Measurement of Surface Topography*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-12012-1.

— (2014). *Fundamental Principles of Engineering Nanometrology*. 2nd. Elsevier.

Lee, S. U., S. Yoon Chung and R. H. Park (1990). "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation". In: *Computer Vision, Graphics, and Image Processing* 52.2, pp. 171–190. DOI: 10.1016/0734-189X(90)90053-X.

Lemieux, C. (2009). *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Verlang.

Leskova, T. A., A. A. Maradudin, E. E. Garcia-Guerrero and E. R. Mendez (2007). "Structured surfaces as optical metamaterials". In: *Metamaterials* 1.1, pp. 19–39. DOI: 10.1016/j.metmat.2007.02.002.

Li, X. Q., Z. W. Zhao, H. D. Cheng, C. M. Huang and R. W. Harris (1994). "A Fuzzy Logic Approach to Image Segmentation". In: *Proceedings of 12th International Conference on Pattern Recognition* 1, pp. 337–341. DOI: 10.1109/ICPR.1994.576291.

Lin, Y.-H. and C.-H. Chen (2008). "Template matching using the parametric template vector with translation, rotation and scale invariance". In: *Pattern Recognition* 41.7, pp. 2413–2421. DOI: 10.1016/j.patcog.2008.01.017.

Lou, S., X. Jiang and P. J. Scott (2013). "Geometric computation theory for morphological filtering on freeform surfaces". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469.2159, p. 20130150. DOI: 10.1098/rspa.2013.0150.

Macqueen, J. (1967). "Some methods for classifcation and analysis of multivariate observations". In: *Fifth Berkley Symposium on Mathematical Statistics and Probability*. Vol. 1, pp. 281–297.

Mainsah, E, J. A. Greenwood and D. G. Chetwynd (2001). *Metrology and Properties of Engineering Surfaces*. Kulwer Academic Publishers.

Malshe, A., K. Rajurkar, A. Samant, H. N. Hansen, S. Bapat and W. Jiang (2013). "Bio-inspired functional surfaces for advanced applications". In: *CIRP Annals - Manufacturing Technology* 62.2, pp. 607–628. DOI: 10.1016/j.cirp.2013.05.008.

Martines, E., K. Seunarine, H. Morgan, N. Gadegaard, C. D. W. Wilkinson and M. O. Riehle (2005). "Superhydrophobicity and Superhydrophilicity of Regular Nanopatterns". In: *Nano letters* 5.10, pp. 2097–103. DOI: 10.1021/nl051435t.

Maxwell, J. C. (1870). "On Hills and Dales". In: *Rep. Meeting Br. Assoc. Adv. Sci.(Liverpool) Trans* 40, pp. 421–425.

Meyer, F. (1994). "Topographic distance and watershed lines". In: *Signal Processing* 38.1, pp. 113–125. DOI: 10.1016/0165-1684(94)90060-4.

Mohamed, N. A., M. N. Ahmed and A. Farag (1999). "Modified Fuzzy C-Mean in Medical Image Segmentation". In: *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 4. IEEE, 3429–3432 vol.6. DOI: 10.1109/ICASSP.1999.757579.

Mountains Map 7. URL: http://www.digitalsurf.com/en/index.html (visited on 09/06/2015).

Nieuwstadt, F. T. M., W. Wolthers, H. Leijdens, K. Krishna Prasad and A. Schwarz-van Manen (1993). "The reduction of skin friction by riblets under the influence of an adverse pressure gradient". In: *Experiments in Fluids* 15.1, pp. 17–26. DOI: 10.1007/BF00195591.

Olympus LEXT OLS4000. URL: http://www.olympus-ims.com/en/metrology/ols4000/ (visited on 11/06/2015).

Otsu, N. (1975). "A Threshold Selection Method from Gray-Level Histograms". In: *Automatica* 9.1, pp. 62–66. DOI: 10.1109/TSMC.1979.4310076.

Pendry, J. B., L. Martin-Moreno and F. J. Garcia-Vidal (2004). "Mimicking Surface Plasmons with Structured Surfaces". In: *Science* 305.5685, pp. 847–848.

Perona, P. and J. Malik (1990). "Scale-space and edge detection using anisotropic diffusion". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.7, pp. 629–639. DOI: 10.1109/34.56205.

Podgornik, B. and M. Sedlacek (2012). "Performance, Characterization and Design of Textured Surfaces". In: *Journal of Tribology* 134.4, p. 041701. DOI: 10.1115/1.4007108.

Raid, I., T. Kusnezowa and J. Seewig (2013). "Application of ordinary kriging for interpolation of micro-structured technical surfaces". In: *Measurement Science and Technology* 24.9, p. 095201. DOI: 10.1088/0957-0233/24/9/095201.

Roach, P., T. Parker, N. Gadegaard and M. R. Alexander (2010). "Surface strategies for control of neuronal cell adhesion: A review". In: *Surface Science Reports* 65.6, pp. 145–173. DOI: 10.1016/j.surfrep.2010.07.001.

Ryk, G., Y. Kligerman, I. Etsion and A. Shinkarenko (2005). "Experimental Investigation of Partial Laser Surface Texturing for Piston-Ring Friction Reduction". In: *Tribology Transactions* 48.4, pp. 583–588. DOI: 10.1080/05698190500313544.

Sahoo, P. K., S. Soltani and A. K. C. Wong (1988). "A Survey of Thresholding Techniques". In: *Computer Vision, Graphics, and Image Processing* 41.2, pp. 233–260. DOI: 10.1016/0734-189X(88)90022-9.

Schwenke, H., B. R. L. Siebert, F. Wäldele and H. Kunzmann (2000). "Assessment of Uncertainties in Dimensional Metrology by Monte Carlo Simulation: Proposal of a Modular and Visual Software". In: *CIRP Annals - Manufacturing Technology* 49.1, pp. 395–398. DOI: 10.1016/S0007-8506(07)62973-4.

Scott, P. J. (2004). "Pattern analysis and metrology: the extraction of stable features from observable measurements". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 460.2050, pp. 2845–2864. DOI: 10.1098/rspa.2004.1291.

Seewig, J. (2005). "Linear and robust Gaussian regression filters". In: *Journal of Physics: Conference Series* 13, pp. 254–257. DOI: 10.1088/1742-6596/13/1/059.

Seewig, J. (2013). "Areal Filtering Methods". In: *Characterisation of Areal Surface Texture*. Ed. by R. K. Leach. Springer. Chap. 4, pp. 67–106.

Senin, N., M. Ziliotti and R. Groppetti (2007). "Three-dimensional surface topography segmentation through clustering". In: *Wear* 262.3-4, pp. 395–410. DOI: 10.1016/j.wear.2006.06.013.

Senin, N. and L. Blunt (2013). "Characterization of individual areal features". In: *Characterisation of Areal Surface Texture*. Ed. by R. K. Leach. Springer. Chap. 8, pp. 179–216.

Senin, N., S. Pini and R. Groppetti (2011). "Identification of Microtopographic Surface Features and Form Error Assessment". In: *Geometric Tolerances: Impact on Product Design, Quality Inspection and Statistical Process Monitoring*. Ed. by B. M. Colosimo and N. Senin. Chap. 5, pp. 159–188.

Senin, N., L. A. Blunt and M. Tolley (2012a). "Dimensional metrology of micro parts by optical three-dimensional profilometry and areal surface topography analysis". In: *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 226.11, pp. 1819–1832. DOI: 10.1177/0954405412461240.

Senin, N., L. Blunt and M. Tolley (2012b). “The use of areal surface topography analysis for the inspection of micro-fabricated thin foil laser targets for ion acceleration”. In: *Measurement Science and Technology* 23.10, p. 105004. DOI: 10.1088/0957-0233/23/10/105004.

Senin, N., L. A. Blunt, R. K. Leach and S. Pini (2013). “Morphologic segmentation algorithms for extracting individual surface features from areal surface topography maps”. In: *Surface Topography: Metrology and Properties* 1.1, p. 015005. DOI: 10.1088/2051-672X/1/1/015005.

Senin, N., G. MacAulay, C. Giusca and R. K. Leach (2014). “On the characterisation of periodic patterns in tessellated surfaces”. In: *Surface Topography: Metrology and Properties* 2.2, p. 025005. DOI: 10.1088/1468-6996/2/2/025005.

Shinkarenko, A, Y Kligerman and I Etsion (2009). “The effect of surface texturing in soft elasto-hydrodynamic lubrication”. In: *International Conference on Tribology*. September.

Singh, A. V. et al. (2012). “Bottom-up engineering of the surface roughness of nanostructured cubic zirconia to control cell adhesion”. In: *Nanotechnology* 23.47, p. 475101. DOI: 10.1088/0957-4484/23/47/475101.

Smith, I. *NPL LSGE library*. URL: http://www.eurometros.org/metros/packages/lsge/ (visited on 11/08/2014).

Song, M. and D. Civco (2004). “Road Extraction Using SVM and Image Segmentation”. In: *Photogrammetric Engineering & Remote Sensing* 70.12, pp. 1365–1371. DOI: 10.14358/PERS.70.12.1365.

Stavroulakis, P. I., S. A. Boden, T. Johnson and D. M. Bagnall (2013). “Suppression of backscattered diffraction from sub-wavelength ’moth-eye’ arrays”. In: *Optics express* 21.1, pp. 1–11. DOI: 10.1364/OE.21.000001.

Stout, K. J. and L. Blunt (2000). *Three Dimensional Surface Topography*. 2nd. Penton press.

— (2001). “A contribution to the debate on surface classificationsâ”random, systematic, unstructured, structured and engineered”. In: *International Journal of Machine Tools and Manufacture* 41.13-14, pp. 2039–2044. DOI: 10.1016/S0890-6955(01)00069-4.

Sun, Y. and G.-j. He (2008). “Segmentation of High-Resolution Remote Sensing Image Based on Marker-Based Watershed Algorithm”. In: *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. IEEE, pp. 271–276. DOI: 10.1109/FSKD.2008.249.

Uehara, Y., M. Wakuda, Y. Yamauchi, S. Kanzaki and S. Sakaguchi (2004). "Tribological properties of dimpled silicon nitride under oil lubrication". In: *Journal of the European Ceramic Society* 24.2, pp. 369–373. DOI: 10.1016/S0955-2219(03)00220-6.

Undeman, C. and T. Lindeberg (2003). "Fully Automatic Segmentation of MRI Brain Images using Probabilistic Anisotropic Diffusion and Multi-Scale Watersheds". In: *Scale Space Methods in Computer Vision*. Vol. 2695, pp. 641–656. DOI: 10.1007/3-540-44935-3_45.

Vincent, L. (1993). "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms." In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 2.2, pp. 176–201. DOI: 10.1109/83.217222.

Vorburger, T. V., R. Silver, R. Brodmann, B. Brodmann and J. Seewig (2011). "Light Scattering Methods". In: *Optical Measurement of Surface Topography*. Ed. by R. K. Leach. Springer. Chap. 12.

Wang, L. (2014). "Use of structured surfaces for friction and wear control on bearing surfaces". In: *Surface Topography: Metrology and Properties* 2.4, p. 043001. DOI: 10.1088/2051-672X/2/4/043001.

Wang, X., K. Kato, K. Adachi and K. Aizawa (2003). "Loads carrying capacity map for the surface texture design of SiC thrust bearing sliding in water". In: *Tribology International* 36.3, pp. 189–197. DOI: 10.1016/S0301-679X(02)00145-7.

Wang, X., K. Adachi, K. Otsuka and K. Kato (2006). "Optimization of the surface texture for silicon carbide sliding in water". In: *Applied Surface Science* 253.3, pp. 1282–1286. DOI: 10.1016/j.apsusc.2006.01.076.

Weckenmann, A. and W. Hartmann (2013). "Function-oriented method for the definition and verification of microstructured surfaces". In: *Precision Engineering* 37.3, pp. 684–693. DOI: 10.1016/j.precisioneng.2013.01.013.

Whitehouse, D. J. (1997). "Surface metrology". In: *Measurement Science and Technology* 8.9, pp. 955–972. DOI: 10.1088/0957-0233/8/9/002.

Whitehouse, D. J. (2010). *Handbook of Surface and Nanometrology*. 2nd. CRC Press.

Wolf, G. W. (1991). "A FORTRAN subroutine for cartographic generalization". In: *Computers & Geosciences* 17.10, pp. 1359–1381. DOI: 10.1016/0098-3004(91)90002-U.

Xiao, S., F. Xie, L. Blunt, P. Scott and X. Jiang (2006). "Feature extraction for structured surface based on surface networks and edge detection". In: *Materials Science in Semiconductor Processing* 9.1-3, pp. 210–214. DOI: 10.1016/j.mssp.2006.01.085.

Xu, R and D Wunsch (2009). *Clustering*. 1st. Wiley-IEEE Press.

Yang, J., F. R. J. Rose, N. Gadegaard and M. R. Alexander (2009). "A High-Throughput Assay of Cell-Surface Interactions using Topographical and Chemical Gradients". In: *Advanced Materials* 21.3, pp. 300–304. DOI: 10.1002/adma.200801942.

Yang, Y. and S. Huang (2012). "Image Segmentation by Fuzzy C-means Clustering algorithm with a novel penalty term". In: *Computing and Informatics* 26.1, pp. 17–31.

Yu, D. P., X. Zhong, Y. S. Wong, G. S. Hong, W. F. Lu and H. L. Cheng (2011a). "An automatic form error evaluation method for characterizing micro-structured surfaces". In: *Measurement Science and Technology* 22.1, p. 015105. DOI: 10.1088/0957-0233/22/1/015105.

Yu, Z., H.-S. Wong and G. Wen (2011b). "A modified support vector machine and its application to image segmentation". In: *Image and Vision Computing* 29.1, pp. 29–40. DOI: 10.1016/j.imavis.2010.08.003.

Zeng, W., X. Jiang, P. J. Scott and T. Li (2013). "A New Method to Characterize the Structured Tessellation Surface". In: *Procedia CIRP* 10, pp. 155–161. DOI: 10.1016/j.procir.2013.08.026.

Zhong, B. and K.-K. Ma (2010). "On the Convergence of Planar Curves Under Smoothing." In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 19.8, pp. 2171–89. DOI: 10.1109/TIP.2010.2046807.

Zhu, H., L. Blunt, X. Jiang and S. Xiao (2009a). "Recognition of Features from Micro Scale Patterned Surfaces". In: *9th International Symposium on Measurement Technology and Intelligent Instruments*.

Zhu, H. (2012). "Measurement and characterisation of micro/nano scale structured surfaces". PhD thesis. Huddersfield.

Zhu, H., L. Blunt and X. Jiang (2009b). "Flexible shape extraction for micro/nano scale structured surfaces". In: *Computing and Engineering Annual Researchers' Conference University of Huddersfield*, pp. 190–194.

Appendix

This appendix provides an overview of the code of the code used in this thesis. Many variations on this code was used and this should not be considered complete code.

This function performs levelling on an image allowing for exclusion of a specified region.

```
%Function for levelling images

function [limg]=level(img, mask)

[x, y]=meshgrid(1:size(img,2), 1:size(img, 1));

xyz = [x(:), y(:), img(:)];        %convert to xyz coords

[yrem, xrem]=find(mask==1);        %find points where mask ==1

tempxyz=xyz;
[tf, mark]=(ismember([xrem, yrem], tempxyz(:,1:2), 'rows')); %remove
    points in mask
mark=[mark; find(isnan(tempxyz(:, 3))==1)];
tempxyz(mark, :)=[];

[x0, a]=lsplane(tempxyz);     %fit lsplane

fitplanez=(a(3)*x0(3)-a(1)*(xyz(:, 1)-x0(1))-a(2)*(xyz(:, 2)-x0(2)))
    /a(3);

fitplane=zeros(size(img));
for i=1:length(xyz)
fitplane(xyz(i,2), xyz(i, 1))=fitplanez(i);
```

```
23 end
24
25 limg=img-fitplane;
26
27 end
```

Function to perform Gaussian filtering with a specified cut-off and not edge correction.

```
1  %Gaussian filtering no edge correction
2
3  function smim = gaussfilt_cut(im, cutoff)
4
5      sigma=cutoff*sqrt(log(2)/2)/pi; %correct filter size so sigma
           gives lambda_c i.e. 50% frequency cutoff.
6
7      if sigma == 0
8          smim = im;
9          return;
10     end
11
12     % If needed convert im to double
13      if ~strcmp(class(im),'double')
14          im = double(im);
15      end
16
17     sze = ceil(6*sigma);
18     if ~mod(sze,2)     % Ensure filter size is odd
19         sze = sze+1;
20     end
21     sze = max(sze,1); % and make sure it is at least 1
22
23     h = fspecial('gaussian', [sze sze], sigma);
24
25     smim = imfilter(im, h, 0);
```

Function to preform Guassian filtering with edge correction.

```
1  %Edge corrected Gaussian filter
2
3  function smim = gaussfiltec(im, cutoff, order)
```

```
4
5       if cutoff == 0
6           smim = im;
7           return;
8       end
9
10      if order~=0
11      x=1:size(im, 2);
12      y=1:size(im, 1);
13      x=repmat(x, size(im, 1), 1);
14      y=repmat(y, 1, size(im, 2));
15
16      x=reshape(x, numel(im), 1);
17      y=reshape(y, numel(im), 1);
18      z=reshape(im, numel(im), 1);
19      x(isnan(z))=[];
20      y(isnan(z))=[];
21      z(isnan(z))=[];
22      end
23
24      if order==0             %correct order if desired this reduces
            edge effects further
25          back=0;
26      elseif order==1
27          backfit=fit([x, y], z, 'poly11');
28          back=feval(backfit, x, y);
29          back=reshape(back, size(im, 1), size(im, 2));
30      elseif order==2
31          backfit=fit([x, y], z, 'poly22');
32          back=feval(backfit, x, y);
33          back=reshape(back, size(im, 1), size(im, 2));
34      end
35
36      im=im-back;
37
38      % If needed convert im to double
39      if ~strcmp(class(im),'double')
```

```
40          im = double(im);
41      end
42
43      sigma=cutoff*sqrt(log(2)/2)/pi; %correct filter size so sigma
            gives lambda_c i.e. 50% frequency cutoff.
44
45      sze = ceil(6*sigma);            %sze needs to be significantly
            bigger than sigma
46      if ~mod(sze,2)     % Ensure filter size is odd
47          sze = sze+1;
48      end
49      sze = max(sze,1); % and make sure it is at least 1
50
51      h = fspecial('gaussian', [sze sze], sigma);
52
53      t_img=imfilter(im, h, 0); %calc filter padding with zeros
54
55      weights=ones(size(im));       %calculate weight to account for
            edge effect
56      lweight=zeros(floor(sze/2));
57
58  for i=1:ceil(sze/2)
59      for j=1:ceil(sze/2)
60          lweight(i, j)=1/sum(sum(h(1:i+ceil(sze/2)-1, 1:j+ceil(sze/2)
                -1)));
61      end
62  end
63  for i=1:size(im, 1)
64      for j=1:size(im, 2)
65          if i<ceil(sze/2) || j<ceil(sze/2) || abs(i-size(im, 1)-1)<
                ceil(sze/2) || abs(j-size(im, 2)-1)<ceil(sze/2)
66      weights(i,j)=lweight(min([i, ceil(sze/2), abs(i-size(im, 1)-1)]),
            min([j, ceil(sze/2), abs(j-size(im, 2)-1)]));
67          end
68      end
69  end
70
```

```
71 smim=t_img.*weights;      %apply weights
72
73      smim=smim+back;      %remove order correction so resembles
            original topo.
```

Demo function to segment a feature and calculate parameters. Variations of this were used to produce the results in chapter 4.

```
1  function [ hparams, gparams, acparams, fimg, hmask, gmask, acmask] =
       segmentation_demo( name, pixsize, filtsize, height_thresh,
       grad_thresh, ac_smooth)
2
3  %Demo to segment an image and calculate parameters using height and
       gradient thresholds and active contours.
4
5  %% load data and calculate gradient image
6  img=read3D(name);    %load image
7
8  fimg=gaussfiltec(img, filtsize/pixsize, 1); %gaussian filter (edge
       corrected)
9
10 gimg=imgradient(fimg);
11 lmask=zeros(size(img));
12
13 %% calculate inital segmentation for levelling
14
15 %here we use gradient threshold can do this using other methods too.
16 if ~isempty(grad_thresh)
17      lmask(gimg>grad_thresh)=1;
18 else
19      lmask(gimg>max(gimg(:))/4)=1;
20 end
21
22 lmask=imfill(lmask, 'holes');      %fill enclosed regions
23 lmask=bwareaopen(lmask, 5000);    %remove small regions (noise/
       roughness)
24 SE=strel('octagon', 12);
25 lmask=imdilate(lmask, SE);     %dilate mask to ensure only background
```

```
    levelled
26
27 fimg=level(fimg, lmask);      %level using background only.
28
29 %% Height thresholding
30
31 if ~isempty(height_thresh)
32     hmask=zeros(size(fimg));
33     hmask(fimg<height_thresh)=1;
34     hmask=imfill(hmask, 'holes');      %fill enclosed regions
35     hmask=bwareaopen(hmask, 5000);    %remove small regions (noise/
           roughness)
36     hmask=imclearborder(hmask);       %remove regions on image edge
37
38     hedg=edge(hmask);      %cal edges of mask
39     edgpoints=[];
40     [edgpoints(:, 2), edgpoints(:, 1)]=find(hedg==1);
41
42     if isempty(edgpoints)
43         warning('no feature detected with height threshold')
44     else
45     [hD, hdepth, hOOR]=calculate_params(edgpoints, fimg, pixsize);
46     hparams=[hD, hdepth, hOOR];
47     end
48 else
49      hparams=[];
50      hmask=[];
51 end
52
53 %%gradient thresholding
54
55 if ~isempty(grad_thresh)
56     gmask=zeros(size(fimg));
57     gmask(gimg>grad_thresh)=1;
58     gmask=imfill(gmask, 'holes');      %fill enclosed regions
59     gmask=bwareaopen(gmask, 5000);    %remove small regions (noise/
           roughness)
```

```
60      gmask=imclearborder(gmask);        %remove regions on image edge
61
62      gedg=edge(gmask);       %cal edges of mask
63      edgpoints=[];
64      [edgpoints(:, 2), edgpoints(:, 1)]=find(gedg==1);
65
66      if isempty(edgpoints)
67          warning('no feature detected with gradient threshold');
68      else
69      [gD, gdepth, gOOR]=calculate_params(edgpoints, fimg, pixsize);
70      gparams=[gD, gdepth, gOOR];
71      end
72      else
73       gparams=[];
74       gmask=[];
75  end
76
77  %% active contours
78
79  if ~isempty(ac_smooth)
80
81  %       acmask=lmask;   %for speed start with good inital guess of
        mask.
82  %       acmask=imclearborder(acmask);       %remove regions on image
        edge (this is not alwys necessary and depnds on how feature is
        defined).
83
84       acmask=zeros(size(fimg));   %alternatively use a box in roughly
             the right place
85       acmask(400:600, 400:600)=1;
86
87       acmask=activecontour(fimg, acmask, 1000, 'Chan-Vese', ac_smooth)
             ;
88       acmask=bwareaopen(acmask, 5000);
89
90       acedg=edge(acmask);       %cal edges of mask
91      edgpoints=[];
```

```
92      [edgpoints(:, 2), edgpoints(:, 1)]=find(acedg==1);
93
94      if isempty(edgpoints)
95          warning('no feature detected with active contours');
96      else
97      [acD, acdepth, acOOR]=calculate_params(edgpoints, fimg, pixsize);
98      acparams=[acD, acdepth, acOOR];
99      end
100     else
101      acparams=[];
102      acmask=[];
103 end
104
105 end
106 %% calculate parameters
107
108 function [D, depth, OOR]=calculate_params(edgpoints, fimg, pixsize);
109 %calculate diameter, depth and out of roundness given edgepoints
110
111 [centre, R, ~, std, ~, ~, uncer]=ls2dcircle(edgpoints, mean(
      edgpoints)', 20, 1e-3, 1e-3); %fit circle to edge
112 D=2*R*pixsize;  %Diameter
113
114   smallcirc=zeros(size(fimg));      %create a small circle with half
        the radius to look at the depth with
115 for j=1:size(smallcirc, 1)
116     for i=1:size(smallcirc, 2)
117         if sqrt((j-centre(2))^2+(i-centre(1))^2)<R/2
118             smallcirc(j, i)=1;
119         end
120     end
121 end
122 depth=-mean(fimg(smallcirc==1));     %depth is mean height within
      circle
123
124      Rpoints=sqrt((edgpoints(:,2)-centre(2)).^2+(edgpoints(:, 1)-
         centre(1)).^2); %radial distance of points
```

```
125
126     OOR=(max(Rpoints)-min(Rpoints))*pixsize;      %out of roundness -
            max-min deviation from circle
127
128 end
```

Function to determine measurement reproducibility from a series of measurements using height thresholding. Variations of this function were used to calculate the results in chapter 5.

```
1  %reproducibility analysis using height threshold
2
3  outd=zeros(250,3);
4  depth=outd;
5  xpos=outd;
6  ypos=outd;
7  rep=outd;
8  cx=outd;
9  cy=outd;
10 rounderr=outd;
11
12 aved=zeros(25,3);
13 avex=aved;
14 avey=aved;
15 repvar=aved;
16 medd=aved;
17 aveoor=aved;
18 varoor=aved;
19
20 pixsize=0.438;        %LEXT 0.625 Alicona 0.438
21
22
23 x_size=1624;          %image size pixels: Alicona 1624x1232, Lext 1024
       x1024
24 y_size=1232;
25
26 %   stepx=75/pixsize; %LEXT
27 %   stepy=75/pixsize;
```

```
28
29 stepx=90/pixsize; %300um
30 stepy=50/pixsize;
31
32 % stepx=130/pixsize; %150um
33 % stepy=90/pixsize;
34
35 %  stepx=165/pixsize; %50um
36 %  stepy=110/pixsize;
37
38
39
40 % [gx, gy]=meshgrid(-2:0.1:2, -2:0.1:2);
41 %
42 % interpr=zeros(size(gx, 1), size(gx, 2), 3);
43 % interpstd=interpr;
44 % interpmedr=interpr;
45 % count=zeros(1000, 1000, 3);
46 for t=1      %:12
47
48 thresh=-0.5;
49
50 for k=1:3
51
52 files=dir(['T:\EM\_MD\Projects\115944 - IRD\Gavin\dimple uncertainty
       \2030010_dimp' num2str(k) '*.al3d']);
53 % files=dir(['T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability
       data Lext 3-6-15\dimp_' num2str(k) '*.lext']);
54 %for n=83
55 for n=1:length(files)
56     disp(n)
57     name=['T:\EM\_MD\Projects\115944 - IRD\Gavin\dimple uncertainty\
           ' files(n).name];
58 %      name=['T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability
       data Lext 3-6-15\' files(n).name];
59
60     matchx=regexp(name, '(?<=x_).*(?=_y)', 'match');
```

```
61      matchy=regexp(name, '(?<=y_).*(?=_rep)', 'match');
62      matchrep=regexp(name, '(?<=rep_).*(?=.al3d)', 'match');
63      xpos(n,k)=str2double(matchx{1});
64      ypos(n,k)=str2double(matchy{1});
65      rep(n,k)=str2double(matchrep{1});
66
67  %       matchcount=regexp(name, '(?<=_000_).*(?=.lext)', 'match');
68  %       matchcount=str2double(matchcount{1});
69  %       xpos(n, k)=mod(mod(matchcount-1, 25), 5)-2;
70  %       ypos(n, k)=(ceil((mod(matchcount-1, 25)+1)/5))-3;
71  %       rep(n, k)=ceil(matchcount/25);
72
73      xest=floor(xpos(n,k)*stepx);
74      yest=floor(ypos(n,k)*stepy);
75
76  %   a=AliconaReader(name);          %to load alicona files instead
77  %     dimple=a.DepthData;
78  %     dimple(dimple==1e10)=NaN;
79  %     dimple=flipud(dimple);
80  %     img=dimple*1e6;
81
82  img=read3D(name);
83
84
85
86      fimg=gaussfilt(img, 4);      %filter image (gaussian sigma=4 => 9
            pixel width)
87      fimg=fimg-mean(fimg(~isnan(fimg)));      %set mean to 0 so initial
            level doesn't fail
88
89      mask=zeros(size(fimg));
90      mask(fimg<thresh)=1;
91
92      mask=imfill(mask, 'holes');      %fill holes
93      fimg=level(fimg, mask);
94
95      mask=zeros(size(fimg));          %recalculate for levelled image
```

```
96      mask(fimg<thresh)=1;
97
98      mask=imfill(mask, 'holes');       %fill holes
99      fimg=level(fimg, mask);
100
101     mask=zeros(size(fimg));           %recalculate for levelled image
102     mask(fimg<thresh)=1;
103
104     mask=imclearborder(mask);
105     mask=imfill(mask, 'holes');       %fill holes
106     %mask=bwareaopen(mask,1000);      %remove small regions. This
            value will need to change depending on size of hole, find a
            better way to do this.
107
108
109     sel=zeros(size(img));
110     sel(ceil(y_size/2)+yest, ceil(x_size/2)-xest)=1;     %Alicona
111 %       sel(ceil(y_size/2)-yest, ceil(x_size/2)+xest)=1;     %Lext
112
113 if ~any(sel==mask & mask==1)
114     mask=bwareaopen(mask,1000);      %remove any small rubbish
115     while ~any(sel==mask & mask==1)
116         sel=bwmorph(sel, 'dilate');  %dilate until they intersect
117     end
118         tsel=sel;
119         sel=zeros(size(img));
120         sel(tsel==1 & mask==1)=1;
121         mask=double(mask);    %some reason makes this logical?
122 end
123
124     tmask=imreconstruct(sel, mask);  %remove all except correct
            dimple
125
126 %       figure
127 %       imagesc(tmask)
128 %       pause(1)
129
```

```
130       edg=edge(tmask);       %cal edges
131       edglist=find(edg==1);
132
133        edgpoints=[];
134       [edgpoints(:,2), edgpoints(:,1)]=ind2sub(size(img),  edglist);
135       [centre, R, ~, stdr, ~, ~, uncer]=ls2dcircle(edgpoints, mean(
              edgpoints)', 20,1e-3,1e-3); %fit circle to edge
136       cx(n,k)=centre(1);
137       cy(n,k)=centre(2);
138       outd(n,k)=2*R*pixsize;     %convert radius to um
139       if outd(n,k, t)==0
140           warning('No feature detected');
141       end
142
143           [y, x]=find(tmask==1);
144       y=y-floor(cy(n, k))+500;
145       x=x-floor(cx(n, k))+500;
146       temp=zeros(1000, 1000);
147       for i=1:length(y)
148       temp(y(i),x(i))=1;
149       end
150  %      count(:, :, k)=count(:, :, k)+temp;
151
152        smallcirc=zeros(size(img));       %create a small circle with
              half the radius to look at the depth with
153  for j=1:size(smallcirc,1)
154       for i=1:size(smallcirc, 2)
155           if sqrt((j-cy(n,k))^2+(i-cx(n,k))^2)<R/2
156               smallcirc(j, i)=1;
157           end
158       end
159  end
160
161  depth(n,k, t)=-mean(fimg(smallcirc==1));       %cal the depth
162
163       Rpoints=sqrt((edgpoints(:,2)-cy(n, k)).^2+(edgpoints(:, 1)-cx(n,
              k)).^2);
```

```
164         Rdiff=(Rpoints-R)*pixsize;
165
166         rounderr(n, k, t)=max(Rdiff)-min(Rdiff);
167
168 %         gimg=imgradient(fimg);
169 %
170 %         figure(1)
171 %         imagesc(tmask)
172 %         figure(2)
173 %         imagesc(fimg)
174 %         colorbar
175 %         figure(3)
176 %         imagesc(gimg)
177 %         colorbar
178 %         pause(1)
179 end
180
181 for i=1:25
182 aved(i,k)=mean(outd((i-1)*10+1:i*10,k));
183 avex(i,k)=xpos((i-1)*10+1,k);
184 avey(i,k)=ypos((i-1)*10+1,k);
185 repvar(i,k)=var(outd((i-1)*10+1:i*10,k));
186 % medd(i,k)=median(outd((i-1)*10+1:i*10,k));
187 end
188
189 end
190 end
191
192 aveD=mean(outd);
193 stdD=std(outd);
194
195 aveoor=mean(rounderr);
196 stdoor=std(rounderr);
197
198 aveddepth=mean(depth);
199 stddepth=std(depth);
```

The following functions were used to calculate the results in chapter 6.

This function determines the contribution of measurement noise.

```
%Calculate the effects of noise on the size of a feature

reps=10;
thresh=-0.5;

clear difdimp difback


for i=1:reps-1;
    disp(i)

%       dat1=read3D(['T:\EM\_MD\Projects\115944 - IRD\Gavin\Lext\repeatabilty test 20% d300 10um\noise test_1_000_00' sprintf('%02d', i), '.lext']);
%       dat2=read3D(['T:\EM\_MD\Projects\115944 - IRD\Gavin\Lext\repeatabilty test 20% d300 10um\noise test_1_000_00' sprintf('%02d', i+1), '.lext']);

dat1=read3D(['T:\EM\_MD\Projects\115944 - IRD\Gavin\rep noise data\dimp_1_000_' sprintf('%02d', i), '.lext']);
dat2=read3D(['T:\EM\_MD\Projects\115944 - IRD\Gavin\rep noise data\dimp_1_000_' sprintf('%02d', i+1), '.lext']);


     dat1=dat1-mean(dat1(:));
     dat2=dat2-mean(dat2(:));

     mask=zeros(size(dat1));
     mask(dat1<thresh)=1;

     mask=imfill(mask, 'holes');       %fill holes
     mask=bwareaopen(mask, 5000);

     ldat1=level(dat1, mask);
     ldat2=level(dat2, mask);

```

```
31      mask=zeros(size(dat1));
32      mask(dat1<thresh)=1;
33
34      mask=imfill(mask, 'holes');       %fill holes
35      mask=bwareaopen(mask, 5000);
36
37      dif=(ldat1-ldat2)/sqrt(2);
38
39      Sqdimp(i)=std(dif(mask==1));
40      Sqback(i)=std(dif(mask==0));
41
42      cSqdimp(i)=sqrt(sum(Sqdimp(1:i).^2)/i);
43      cSqback(i)=sqrt(sum(Sqback(1:i).^2)/i);
44
45      if exist('difdimp', 'var')
46      difdimp=[difdimp; dif(mask==1)];
47      else
48          difdimp=dif(mask==1);
49      end
50         if exist('difback', 'var')
51      difback=[difback; dif(mask==0)];
52      else
53          difback=dif(mask==0);
54      end
55  end
56
57  dimp=read3D('T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability
        data Lext 3-6-15\dimp_1_000_0013.lext');
58
59   dimp=dimp-mean(dat1(:));
60
61      mask=zeros(size(dimp));
62      mask(dat1<thresh)=1;
63
64      mask=imfill(mask, 'holes');       %fill holes
65      mask=bwareaopen(mask, 5000);
66
```

```
67      dimp=level(dimp, mask);
68
69      mask=zeros(size(dimp));
70      mask(dat1<thresh)=1;
71
72      mask=imfill(mask, 'holes');        %fill holes
73      mask=bwareaopen(mask, 5000);
74
75  for i=1:5000
76      disp(i)
77
78      nmap=rand(size(dimp));
79      nback=difback(ceil(nmap*length(difback)));
80      ndimp=difdimp(ceil(nmap*length(difdimp)));
81
82      n=nback;
83      n(mask==1)=ndimp(mask==1);
84
85      n2=zeros(size(nmap));
86      n2(mask==1)=(nmap(mask==1)-mean(nmap(mask==1)))*cSqdimp(end);
87      n2(mask==0)=(nmap(mask==0)-mean(nmap(mask==0)))*cSqback(end);
88
89      noisedimp=dimp+n;
90      noisedimp2=dimp+n2;
91
92
93      [~, ~, ~, ~, depth(i), ~, ~, R(i)]=run_heightdimp_local(
            noisedimp, 0.625, -0.5);
94      [~, ~, ~, ~, depth2(i), ~, ~, R2(i)]=run_heightdimp_local(
            noisedimp2, 0.625, -0.5);
95  end
96
97  [~, ~, ~, ~, depth0, ~, ~, R0]=run_heightdimp_local(dimp, 0.625,
        -0.5);
98
99  depthbias=mean(depth)-depth0;
100 D=2*R;
```

```
101  Dbias=mean(D)-2*R0;
102  D=sort(D);
103  depth=sort(depth);
104
105 ubdepth=depth(ceil(0.975*5000));
106 ubD=D(ceil(0.975*5000));
107 lbdepth=depth(ceil(0.025*5000));
108 lbD=D(ceil(0.025*5000));
```

This function calculates the flatness map from a series of measurements.

```
1  function [aveimg]=flatness_ave_img(path, textout)
2
3  %may run out of memory for more than ~10-20 images
4  %will fail if you have any non-data files in your folder.
5  %and yes this puts a .txt folder there so move or delete it before
6  %re-running the script.
7
8  %textout is flag to output text file set to 1 to output.
9
10 files=dir([path '\*.*']);
11
12 files=files(~[files.isdir]);
13
14 for i=1:length(files)
15     disp(i)
16
17     z=read3D([path '\' files(i).name]);
18
19     z=level(z, zeros(size(z)));
20
21     z(:, 1)=z(:, 2);     %very tilted flats are for some reason
           losing first column,
22
23     [x, y]=meshgrid(1:size(z, 2), 1:size(z, 1));
24
25     x=reshape(x, numel(x), 1);
26     y=reshape(y, numel(y), 1);
27     tz=reshape(z, numel(z), 1);
```

```
28      f=fit([x(1:10:end),y(1:10:end)], tz(1:10:end), 'poly55', '
            normalize','on');  %remove polynomial
29      ev=feval(f, [x, y]);
30      ev=reshape(ev, size(z));
31  %      ev=gaussfiltec(z, 2/0.625, 1);  %or filter (curretly put in
        pixel size manually)
32      temp=z-ev;
33
34      Sq_temp=std(temp(:));
35
36      z(temp>mean(temp(:))+3*Sq_temp | temp<mean(temp(:))-3*Sq_temp)=
            nan;
37      %z=grubbs(temp,3); %this seems to remove too much?
38
39      mask=zeros(size(z));
40      mask(isnan(z))=1;
41
42      z=level(z, mask);
43
44      if exist('sumdat', 'var')
45          a(:, :, 1)=sumdat;
46          a(:, :, 2)=z;
47          sumdat=nansum(a, 3);
48      else
49          sumdat=z;
50      end
51      aveimg=sumdat/i;
52
53  %          imagesc(z)
54  %          pause(0.5)
55
56
57  end
58
59  for i=1:length(files)
60      disp(i)
61
```

```
62      z=read3D([path '\' files(i).name]);
63
64      z=level(z, zeros(size(z)));
65
66      z(:, 1)=z(:, 2);     %very tilted flats are for some reason
            losing first column,
67
68      [x, y]=meshgrid(1:size(z, 2), 1:size(z, 1));
69
70      x=reshape(x, numel(x), 1);
71      y=reshape(y, numel(y), 1);
72      tz=reshape(z, numel(z), 1);
73      f=fit([x(1:10:end),y(1:10:end)], tz(1:10:end), 'poly55', '
            normalize','on');  %remove polynomial
74      ev=feval(f, [x, y]);
75      ev=reshape(ev, size(z));
76  %     ev=gaussfiltec(z, 2/0.625, 1);  %or filter (curretly put in
        pixel size manually)
77      temp=z-ev;
78
79      Sq_temp=std(temp(:));
80
81      z(temp>mean(temp(:))+3*Sq_temp | temp<mean(temp(:))-3*Sq_temp)=
            nan;
82      %z=grubbs(temp,3); %this seems to remove too much?
83
84      mask=zeros(size(z));
85      mask(isnan(z))=1;
86
87      z=level(z, mask);
88
89      if exist('sumdat', 'var')
90          a(:, :, 1)=sumdat;
91          a(:, :, 2)=z;
92          sumdat=nansum(a, 3);
93      else
94          sumdat=z;
```

```
95      end
96      aveimg=sumdat/i;
97
98  %       imagesc(z)
99  %       pause(0.5)
100
101
102 end
103
104 %aveimg=gaussfiltec(aveimg, 20, 2);      %remove short wavelength
        components
105 %of flatness
106
107 if textout==1
108 dlmwrite([path '\aveimg.txt'], aveimg);
109 end
```

This function calculates the effect of flatness on a measurement from the flatness map.

```
1  %calcualte the effect of the flatness deviation on the size of a
       feature
2
3  img=lextheight('T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability
       data Lext 3-6-15\dimp_1_000_0013.lext');
4
5  flatmap=dlmread('T:\EM\_MD\Projects\115944 - IRD\Gavin\Flatness Lext
        13-3-15\0mm\20x\aveimg.txt');
6
7  rflatimg=img-flatmap;
8
9  x=dlmread('data\corxshift.txt');
10 y=dlmread('data\coryshift.txt');
11
12 modr=zeros(1, 1);
13 moddepth=zeros(1, 1);
14
15 for i=1:5000
16
17     disp(i)
```

```
18      shiftflat=circshift(flatmap, [y(i), x(i)]);
19
20      if x(i)>0
21          shiftflat(:, 1:x(i))=0;
22      elseif x(i)<0
23          shiftflat(:, end+x(i)+1:end)=0;
24      end
25
26          if y(i)>0
27          shiftflat(1:y(i), :)=0;
28      elseif y(i)<0
29          shiftflat(end+y(i)+1:end, :)=0;
30          end
31
32      modimg=rflatimg+shiftflat;
33
34      [ ~, modr(i), fimg, mask, moddepth(i), ~, ~, ~]=
            run_heightdimp_local( modimg, 0.625, -0.5);
35
36  end
37
38  [ ~, r0, fimg, mask, depth0, ~, ~, ~]=run_heightdimp_local( rflatimg
        , 0.625, -0.5);
39
40  D=2*modr;
41  depthbias=mean(moddepth)-depth0;
42  Dbias=mean(D)-2*r0;
43  depth=sort(moddepth);
44  D=sort(D);
45  ubdepth=depth(ceil(0.975*5000));
46  ubD=D(ceil(0.975*5000));
47  lbdepth=depth(ceil(0.025*5000));
48  lbD=D(ceil(0.025*5000));
49
50  save('data\flatness_effect_data_cor');
```

This function calculates the effect of lateral distortion from the lateral distortion map.

```
1  %calculate effect of lateral scale distortion on the size of a
```

```
    feature
2
3 img=lextheight('T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability
    data Lext 3-6-15\dimp_1_000_0013.lext');
4
5 errormapX=dlmread('data\lat_errors_X.txt');
6 errormapY=dlmread('data\lat_errors_Y.txt');
7
8
9 x=randi(300, 5000, 1)-150;
10 y=randi(300, 5000, 1)-150;
11
12 reps=1;
13
14  R=zeros(reps, 5000);
15
16 [img, fimg, mask, edg]=run_heightdimp_local_edge( img, 0.625, -0.5);
17 coredg=zeros(size(edg, 1), 2);
18
19 for i=1:size(edg, 1)
20 coredg(i, 1)=edg(i,1)-errormapX(edg(i, 2), edg(i, 1));
21 coredg(i, 2)=edg(i, 2)-errormapY(edg(i, 2), edg(i, 1));
22 end
23
24  [c0, R0, d0, sigmah0, conv0, Vx0n0, urn0, GNlog0, a0, ~] =
    ls2dcircle(coredg, mean(coredg)', 150, 1e-3, 1e-3);
25
26     smallcirc=zeros(size(img));      %create a small circle with half
        the radius to look at the depth with
27 for n=1:size(smallcirc, 1)
28     for m=1:size(smallcirc, 2)
29         if sqrt((m-c0(2))^2+(n-c0(1))^2)<R0/2
30             smallcirc(n, m)=1;
31         end
32     end
33 end
34
```

```
35 depth0=-mean(fimg(smallcirc==1));      %cal the depth
36
37  R0=R0*0.625;
38
39 for i=1:5000
40      disp(i)
41
42     shifterrorX=circshift(errormapX, int16([y(i), x(i)]));
43     shifterrorY=circshift(errormapY, int16([y(i), x(i)]));
44
45     if x(i)>0
46         shifterrorX(:, 1:x(i))=NaN;
47         shifterrorY(:, 1:x(i))=NaN;
48     elseif x(i)<0
49         shifterrorX(:, end-x(i)+1:end)=NaN;
50         shifterrorY(:, end-x(i)+1:end)=NaN;
51     end
52
53         if y(i)>0
54         shifterrorX(1:y(i), :)=NaN;
55         shifterrorY(1:y(i), :)=NaN;
56     elseif y(i)<0
57         shifterrorX(end-y(i)+1:end, :)=NaN;
58         shifterrorY(end-y(i)+1:end, :)=NaN;
59        end
60
61     shiftedg=zeros(size(edg, 1), 2);
62         for j=1:size(edg, 1)
63     shiftedg(j, 1)=coredg(j, 1)+shifterrorX(edg(j, 1), edg(j, 2));
64     shiftedg(j, 2)=coredg(j, 2)+shifterrorY(edg(j, 1), edg(j, 2));
65         end
66
67  for k=1:reps
68     %shiftedg=shiftedg+randn(size(shiftedg))*0.3; %add random
          uncertainty (in pixels).
69
70    [c(i, :), R(k, i), d, sigmah, conv, Vx0n, urn, GNlog, a, ~] =
```

```
        ls2dcircle(shiftedg, mean(shiftedg)', 150, 1e-3, 1e-3);
71
72      smallcirc=zeros(size(img));      %create a small circle with half
            the radius to look at the depth with
73  for n=1:size(smallcirc, 1)
74      for m=1:size(smallcirc, 2)
75          if sqrt((m-c(i, 2))^2+(n-c(i, 1))^2)<R(k, i)/2
76              smallcirc(n, m)=1;
77          end
78      end
79  end
80
81  depth(k, i)=-mean(fimg(smallcirc==1));      %cal the depth
82
83  end
84  end
85
86  R=R*0.625;
87  dif=R-R0;
88
89  depth=sort(depth);
90  D=2*R;
91  D=sort(D);
92  ubdepth=depth(ceil(0.975*5000));
93  ubD=D(ceil(0.975*5000));
94  lbdepth=depth(ceil(0.025*5000));
95  lbD=D(ceil(0.025*5000));
96
97  depthbias=mean(depth)-depth0;
98  Dbias=mean(D)-2*R0;
```

This function estimates the effect of vertical distortions on the measurement.

```
1  %calculate the effects of the vertical scales on the size of a
       feature
2
3  heights=xlsread('stepcomb20x.xlsx');
4  nomstep=xlsread('calibrated step height values.xlsx');
5
```

```
for i=1:size(heights, 1);

    k=find(nomstep(:, 1)==heights(i, 2));

    heights(i, 8)=nomstep(k, 2);

end

I=find(heights(:, 2)==19);          %exclude 19mm for now due to
    flatness
heights(I, :)=[];

heights(:, 9)=heights(:, 7)-heights(:, 8);

I=find(heights(:, 1)==20);
h20=heights(I, :);
for j=2:7
    I=find(heights(:, 2)==nomstep(j, 1));
    T=heights(I, :);
    mh20(j-1, :)=mean(T);
    sdh20(j-1)=std(T(:, 7))/sqrt(size(T, 1));
end

[f20, gof, flag]=fit(mh20(:, 8), mh20(:, 7), 'poly1', 'lower', [-inf
    , 0], 'upper', [inf, 0]);

amp20=f20.p1;

res20=h20(:, 7)-h20(:, 8)*amp20;

nomres20=h20(:, 7)-h20(:, 8);

%assess uncertainty by MC

for m=1:5000
    disp(m)

```

```
41      Tnomstep=nomstep;
42      Tnomstep(:, 2)=Tnomstep(:, 2)+randn(size(Tnomstep, 1), 1).*
            Tnomstep(:, 3)/2;
43
44      Tmh20=mh20;
45
46      Tmh20(:, 7)=Tmh20(:, 7)+(sdh20.*randn(size(sdh20)))';
47
48      for i=1:size(Tmh20, 1);
49
50      k=find(Tnomstep(:, 1)==Tmh20(i, 2));
51
52      Tmh20(i, 8)=Tnomstep(k, 2);
53      end
54
55      Tf20=fit(Tmh20(:, 8), Tmh20(:, 7), 'poly1', 'lower', [-inf, 0],
            'upper', [inf, 0]);
56
57      Tamp20(m)=Tf20.p1;
58
59  end
60
61  nonlin=max(abs(res20));
62  ampuncert=2*std(Tamp20);
63  mamp=mean(Tamp20);
64
65  img=read3D('T:\EM\_MD\Projects\115944 - IRD\Gavin\repeatability data
        Lext 3-6-15\dimp_1_000_0013.lext');
66
67  for i=1:5000
68      disp(i)
69
70      amp(i)=mamp+randn()*ampuncert+(rand()-0.5)*nonlin/10000;
71
72      limg=img*amp(i);
73
74      [~, R(i), fimg, mask, depth(i), ~, ~, ~]=run_heightdimp_local(
```

```
        limg, 0.625, -0.5);
end

D=2*R;

ubdepth=depth(ceil(0.975*5000));
ubD=D(ceil(0.975*5000));
lbdepth=depth(ceil(0.025*5000));
lbD=D(ceil(0.025*5000));

[ ~, R0, fimg, mask, depth0, ~, ~, ~]=run_heightdimp_local( img,
    0.625, -0.5);
Dbias=mean(D)-2*R0;
depthbias=mean(depth)-depth0;
```

Finally, the combined effect of all metrological characteristics is determined.

```
%Calculate the combined uncertainty from the individual effects

clear all
load('flatness_effect_data3.mat')

c1=D;    %flatness
d1=moddepth;

load('noise_effect3.mat')

c2=D;    %noise
d2=depth;

load('lat_scale_effect 22-7-15.mat')

c3=D;    %lat scales
d3=depth;

load('vert_scale_data3.mat')

c4=D;    %vert scales
d4=depth;
```

```
23
24 rD=ceil(rand(50000, 4)*5000);
25 rdepth=ceil(rand(50000, 4)*5000);
26
27 filter_std=13;              %filter and lat scales in nm
28 rep_latscales_std=58;
29
30 filter_eff=(rand(1, 50000)-0.5)*filter_std*sqrt(12)/1000;
31 rep_lat_eff=randn(1, 50000)*rep_latscales_std/1000;
32
33 bias_minmax_D=194;          %uncertainty in bias in nm
34 bias_minmax_depth=5;
35 bias_std_D=99;
36 bias_std_depth=2.5;
37
38 r1=rand(1, 50000);
39 r2=rand(1, 50000);
40
41 as_D=-(bias_minmax_D/2)-bias_std_D+2*bias_std_D*r1;
42 as_depth=-bias_minmax_depth/2-bias_std_depth+2*bias_std_depth*r1;
43 bs_D=-as_D;
44 bs_depth=-as_depth;
45
46 bias_uncert_D=(as_D+(bs_D-as_D).*r2)/1000;
47 bias_uncert_depth=(as_depth+(bs_depth-as_depth).*r2)/1000;
48
49 %uncorrelated
50  D=c1(rD(:, 1))-mean(c1)+c2(rD(:, 2))-mean(c2)+c3(rD(:, 3))-mean(c3)
       +c4(rD(:, 4))-mean(c4)+filter_eff+rep_lat_eff+bias_uncert_D+(
       mean(c1)+mean(c2)+mean(c3)+mean(c4))/4;
51  depth=d1(rdepth(:, 1))-mean(d1)+d2(rdepth(:, 2))-mean(d2)+d3(rdepth
       (:,3))-mean(d3)+d4(rdepth(:, 4))-mean(d4)+bias_uncert_depth+(
       mean(d1)+mean(d2)+mean(d3)+mean(d4))/4;
52
53 %flatness & lat dist correlated
54 % D=c1(rD(:, 1))-mean(c1)+c2(rD(:, 2))-mean(c2)+c3(rD(:, 1))-mean(c3
       )+c4(rD(:, 3))-mean(c4)+filter_eff+rep_lat_eff+bias_uncert_D+(
```

```
    mean(c1)+mean(c2)+mean(c3)+mean(c4))/4;
% depth=d1(rdepth(:, 1))-mean(d1)+d2(rdepth(:, 2))-mean(d2)+d3(
    rdepth(:,1))-mean(d3)+d4(rdepth(:, 3))-mean(d4)+bias_uncert_depth
    +(mean(d1)+mean(d2)+mean(d3)+mean(d4))/4;

D=sort(D);
depth=sort(depth);

topcut=ceil(0.975*50000);
botcut=ceil(0.025*50000);

uncert_top_D=D(topcut);
uncert_top_depth=depth(topcut);

uncert_bot_D=D(botcut);
uncert_bot_depth=depth(botcut);

mD=mean(D);
sdD=std(D);

mdepth=mean(depth);
sddepth=std(depth);
```